BULLETIN OF THE EXTENSION
DIVISION, INDIANA UNIVERSITY

Vol. VI BLOOMINGTON, IND. No. 12

LIBRARY
APR 5 1922
UNIVERSITY OF TORONTO

Eighth
Conference on Educational Measurements

AUGUST, 1921

Entered as second-class mail matter, October 15, 1915, at the post-office at Bloomington, Indiana, under the Act of August 24, 1912. Published monthly, by Indiana University, from the University Office, Bloomington, Indiana.

For Sale by the University Bookstore, Bloomington, Ind.
Price 50 cents

A limited number of copies of this bulletin will be distributed free of charge to citizens of Indiana.

EIGHTH ANNUAL CONFERENCE

ON

EDUCATIONAL MEASUREMENTS

Held at Indiana University, Bloomington, Ind., Friday and Saturday, April 22 and 23, 1921

PUBLISHED BY
THE EXTENSION DIVISION OF INDIANA UNIVERSITY
1921

Contents

ADDRESS OF WELCOME. By WILLIAM LOWE BRYAN, President of Indiana University 3

CONTRIBUTION OF MEASUREMENTS TO METHODS OF TEACHING. By S. A. COURTIS, Director of the Department of Educational Research, Detroit Public Schools 5

THE RATING SCALE SYSTEM IN DETROIT. By S. A. COURTIS 12

ANALYSIS OF SCHOOL SUBJECTS (Abstract). By CHARLES H. JUDD, Director of the School of Education, University of Chicago 15

PERIODS OF EDUCATIONAL DEVELOPMENT (Abstract). By CHARLES H. JUDD 16

AN EFFICIENT METHOD FOR MEASURING THE RESULTS OF INSTRUCTION IN COLLEGES AND SCHOOLS. By WILLIAM F. BOOK, Head of Department of Psychology and Philosophy, Indiana University 18

MEASUREMENT OF PROGRESS IN ENGLISH IN THE UPPER GRADES. By SIDNEY L. PRESSEY, Department of Psychology and Philosophy, Indiana University 35

READING SCALES FOR THE SECOND, THIRD, AND FOURTH GRADES. By LUELLA W. PRESSEY, Research Fellow in Psychology, Indiana University 46

STANDARDIZED MEASUREMENTS AS A MEANS OF DETERMINING MOST EFFECTIVE CLASS SIZE. By HUBERT G. CHILDS, Professor of Secondary Education, Indiana University 53

DIAGNOSTIC TEST IN MODERN EUROPEAN HISTORY. By CHARLES G. VANNEST, Critic Teacher in History, Indiana University 59

SCHOOL BOARDS IN INDIANA. By WILLIS A. HOLIMAN, Superintendent of Schools, Martinsville, Ind 69

Address of Welcome

WILLIAM LOWE BRYAN

IT is greatly to be desired that the efforts to discover a persistence test should be successful. An adequate persistence test would be of more value than an intelligence test. Individuals are lifted out of their intelligence class up or down by their greater or less persistence. Until such test can be made the fact must be taken account of as well as possible by informal means.

In every science laboratory experimentation should so far as possible be supplemented by field observation. In the laboratory we eliminate disturbing conditions in order to secure precision, but in practice these disturbing conditions are again present and may be so important as to quite upset the results of the laboratory studies. In the best modern biological studies the two methods are united. They should by all means be united in the study of child life.

Digitized by the Internet Archive
in 2008 with funding from
Microsoft Corporation

http://www.archive.org/details/no12conferenceoned6indiuoft

Contributions of Measurements to Methods of Teaching

S. A. Courtis

I WANT to speak tonight on the topic I am most interested in—methods of teaching. The title, "Contributions of Measurements to Methods of Teaching", gives me the chance to discuss an idea about measurement work that some of us have had from the beginning, altho this point of view has not always been understood by some of the teachers who have used our tests.

As I see it, tests are a means of controlling methods of teaching. They are of no value except to the teacher who uses the results intelligently. I want to try to show you tonight in various ways how the use of tests has contributed to the building up of a knowledge of the proper ways of teaching, and how, as we work with tests in the different fields, we are constructing a perfectly definite philosophy of teaching which is enabling us to do better and better work.

I am going to begin with a definition of education in order that we may start with the same basic ideas. This is the definition of education which I like best: Education is a process of helping children to change themselves from where they are toward where we think they ought to be. I like that definition for many reasons. First, I like to emphasize the fact that teaching is a process in which a number of different people have to coöperate. Education is not a matter of taking a course here and there. It is something more than that. What you can accomplish is determined in large measure by what was done by those who came before you, while what you do will determine in part the success or failure of those who come after you. Coöperation is essential to success.

Then education ought to be a process of helping children. That puts the emphasis where it belongs. Children must change themselves; the teacher cannot do it for them. It is the teacher's place *to help* the children change themselves. Now the children are going to change themselves from where they are. The only kind of teaching that counts is that in which we begin to work with the child where he is. We are going to change him from where he is toward the place where we think he ought to be. The tests help us to measure where the children are; and when we know where the children are and where they ought to be, the tests can help us to decide whether a new method is an improvement over the one we have had.

The definition is of value because it enables us to show clearly the functions of measurements. They are: (1) the definition of goals; (2) the description of children; (3) the determination of the efficiency of different methods. With the aid of measurements we have been defining the goal of education. For instance, in the case of handwriting, measurement has helped us to define the goal. The standard was set up by Ayres, after the Thorndike scale had given us some idea how to go to work to measure handwriting. Writing ability is measured in two dimensions, rate and quality. When we find a second grade child who writes at the rate of 32 letters a minute with a quality of 38, we have an illustration of the average of third grade ability. If a man 80 years old writes at this rate and quality, he is third grade in writing ability. In other words, grade becomes a level of ability instead of a group of children. When we measure children in respect to such standards, we can see where they stand.

I do not think people half realize the contribution measurements has made to definiteness of aim. For instance, some people say that we do not write as well in school today as children used to write. We find some people, who have not had measurement work, saying they would like to go back to the good old times. This is a reasonable conclusion if you reach it on the basis of their facts. That is quite easy to prove.

We found in the city of Detroit, in the attic of our board of education building, a book containing specimens of writing of 500 children prepared for the centennial exhibition in 1876. There were 500 samples from the second to the eighth grade, and every one of them in copper plate writing. Anyone who remembers the writing of those days would be inclined to say: "These new-fangled ways of teaching are no good. When I went to school our teacher would not let us turn in a paper with writing which was not perfectly made. We had to produce perfect copies." When our writing supervisor found this book, a question immediately suggested itself to all of us: "What is the writing of those people like today?" From 1876 to the present time makes about time enough for those people to have been out of school forty-five years. For instance, there was a sample in this book from a boy who is now principal of one of our largest high schools, and another from a girl who is now supervisor of English. Altogether we were able to locate about 90 people in Detroit whose writing as children was in that book of samples. So we sent to these people and asked them to write the same material, then compared it with their writing today. The 90 samples average about 55 Ayres, and many of the samples were very poor indeed. Was that writing in 1876 so fine then, or did they have wrong standards? They succeeded in achieving the goals they aimed at. All the writing done was copper plate writing, but how about the method by which it was achieved? The men and women who wrote for us told about the slow, painful process by which the fine specimens of 1876 were brought about. Under this training they acquired a kind of ability in writing which was not at all the same ability to write that is desired in life.

We know today that the writing we teach in school must be of such character that it will stand up under ordinary usage. We must consider rate as well as quality. The standard we set today for the eighth grade is quality 60 and a rate of about 85 to 90 letters a minute, which insure the child writing habits that will produce legible writing under conditions of actual life use. Many investigations have shown that quality 60 is about the quality needed in various commercial situations where writing is used. We have evidence that a quality of about 60, with a rate of 85 to 90 or better, is a reasonable goal for schools today. In practically every sort of educational work there is going on a similar definition of the objectives toward which we work.

So much for definition of goal. What about measurement of children?

We base the teaching of children today on measurement of their actual abilities. This was not true of the days that have gone.

One of our high school teachers of English collected a number of samples of English composition, some written by children, some by authors of well known reputation, copied in children's writing. He had samples ranging all the way from very fine compositions to very poor qualities. He submitted them to the teachers in Detroit high school and in high schools in other places. He asked the teachers to pick out of this collection of compositions that sample which was the very lowest they would accept for entrance to high school. That quality of composition, then, would be set up as a minimum for a child to obtain for entrance to high school. In this way he got from the English teachers a statement of their goals. After they had selected the compositions he measured them by the Hillegas scale by which we can interpret merit in composition in absolute terms. At the same time he measured the actual composition ability of the entering class. The teacher's value was about 75 on the Hillegas Scale and the pupil's value about 50. This represents four years of average progress. That is, the teachers tended to set standards three and four years above the actual abilities of the children, a goal probably impossible of achievement. This same situation exists in arithmetic, in history, and all along the line. In teaching it is impossible to plan our work intelligently unless first of all we have an objective standard, then use it to measure the children and see where they stand, then base our work on their actual needs.

What contribution has education made to method? The contributions have been of many types and characters. I am going to take one of the most recent studies in reading from the University of Chicago as an illustration.

A study was made of the voice- and eye-span in reading. They recorded and measured the movements of the eye and kept track also of where the voice was at the same time. The eye and voice moved independently. When the eye was a certain distance in advance, the voice began to read. Usually the voice was several words behind the eye, but at the end of the sentence there was a long pause while the voice caught up. As we study such a record we come to understand the play of the eye and voice together. In other words, these exact studies of what goes on in school work show that reading is a very complex process, and that the child reads, not by looking at each letter or word, but by pausing at certain points and taking in a whole area at a time. This proves very clearly that the old method of teaching by letters was wrong.

We changed from the letter method to the word and sentence method sometime ago. But we have in many schools today intensive drills on methods of making sounds, so that the child may be able to pronounce the words he reads. The people who advocate these methods claim that it is quite impossible for a child to learn except by this system. However, we older people were not taught that system when we went to school, and we certainly cannot use it now. We would have to be instructed before we could use it. Yet somehow we learned to read. Some day somebody is going to come along and do for the oral reading of children precisely the same thing that has been done for the eye- and voice-span. Our present phonic methods are probably wrong. From the beginning we must put children in situations were they may establish as early as possible those habits they will use thruout life. The problem is clearly defined, is it not? For the beginning reading we must adopt some scheme which will enable the child to work in the right way from the very first. Measurement is helping us to develop that scheme.

We are trying to set up in Detroit a method which will enable children to help themselves. We are trying to develop a series of self-help methods in the teaching of reading. We believe that we must teach the child to help himself. Here is a teaching sheet with two words, *apple* and *tree*. He is given a special dictionary with the words he must use, like *green* and *red*, illustrated. He learns, by demonstration with the teacher, to cut out and paste and color. He cuts out the tree, colors it, and he pastes the apple on the tree, being guided in his making of the picture by the printed words. He learns the words thru this construction of a simple picture. Then as a test he is given a little story about the picture to read. When he finishes his constructive work he compares it with the model on the teacher's desk, and sees whether it is correct. If it is correct, he is allowed to try to read. When the test proves he has mastered the story, he can take it home and show his parents. Thus he gets a chance to see directly the reading process as a process by which he accomplishes certain results that he wishes to achieve. In other words, he learns to use his reading as a tool from the beginning.

Now what happens when we try this with children? They learn to read at a most amazing rate. This method produces self-activity, and the value of the results in terms of self-direction, self-appraisal, and self-control is much greater than in the ordinary reading situation. Some teachers say, "All *that* machinery to learn two or three words?" Wait a minute! What is education for? Is reading the final outcome of education for which you work? Or is reading merely a by-product? Reading is merely a means to an end, a tool. We ought to work for power, not merely skill. The new method yields larger returns in both power and personality. Let me show you, if I can, what I mean.

Let us take arithmetic. Here are some conventional textbook problems. We are going to call this one the star-apple problem.

Fig. I. (Omitted by the author.)

Here is the same problem presented in a different way.

Figure II. (Omitted by the author.)

The problem is not presented in words, but in a situation. The child is confronted with a situation in which he is called upon to make adjustments. He must use his arithmetic as a means to an end, as a tool. Can children do that? You are working to produce changes in the child that will be of real value in life, are you not? We have the idea in Detroit that such things as reading and writing are not the real essentials, that there are some other things going on of more importance.

Let me show you what I mean. A comparison of the results from these two tests with a class in the University of Chicago last summer and an eighth grade class in Detroit shows that two results are practically the same. The children would do the textbook problems as well as adults. The score for both was 95 per cent correct. But with the situation, the score for the children was but 42 per cent correct, while for the adults it was 88 per cent.

What is it that the business man wants of his employees? What is it we are teaching our children in school,—to follow certain instructions that we give? Or do we want them intelligent, so they can take hold of a situation and make adjustments in order to achieve a given end? Isn't it barely possible that we have missed the most vital elements in education by choosing the wrong goals? Measurement will help us to find out. In Detroit we believe that the reason why children can do problems in words and cannot handle actual situations is because we have not been working to develop the right kind of abilities. We have been training for arithmetic and not for the development of the children. They fail because they do not have the initiative to tell themselves what to do. They do not have the practical ability to analyze the situation and put the knowledge which they have to practical use.

Let me tell you the story of a boy. One of his friends had a rifle given to him by his father. Several of the boys in the neighborhood had rifles, and this boy wanted a rifle too. His father being wise, said, "Ask your mother." His mother said, "No, city boys should not have rifles"; that she did not believe in rifles. The boy went back to his father and said, "Didn't you have a rifle when you were a boy, and didn't you enjoy it?" The father acknowledged that he did, but he could not go back on the mother, nor against the whole family who also objected to the rifle. So he said, "If you will earn money enough to buy a rifle, I will let you have it." The father thought, of course, that that would be the end of the rifle, but the boy made good.

Now what happened in the situation? The boy did not like English composition in school, but he cared enough about that rifle to write letters to all the firms that had rifles to sell. He can tell you the price and the bore and the kind of ammunition for every rifle; in fact, he can tell you all about rifles. After he got the rifle, what do you suppose he did? He found a young man who had just returned from France, and he pestered the life out of that soldier to drill him in shooting. He actually begged to be taught. Couldn't we get that same kind of situation in school? Aren't the qualities that were developed in the boy, the initiative, drive, and executive ability, those that we need in everyday life? How far are our schools at present organized so that we are working to develop those abilities in all the work that we do?

We were told in the last war that it was a conflict between two kinds of government, autocracy and democracy. We were told that in an autocracy the common people needed only two qualities, obedience and technical skill. If they have those qualities and do as they are told, the autocracy gets along well. But is that the kind of training needed in a democracy? Some of us say not; that the important thing in a democracy is, first of all, for the child to get from his school training an ideal of service as a basis for the organization of all his life purposes. This country is an individualistic country. We believe in the development of the individual, but if we develop the individual too far, we tend towards anarchy. On the other hand, if we develop the individual too little, we drift into communism or some other form of socialism. It is quite evident that we need to keep in the middle of the road. We have to respect the nature of the individual and fit our education around him; and then in order to offset individual tendencies we have to make him so conscious of the contributions the past has made to him, and of the contributions which he can make to his fellow-men, that all his life long he will be doing his best to fulfil his obligation to the past. If I were a supervisor coming into your room, that is the first basis upon which I would judge your work. Are you handling each lesson in such a fashion that you bring out the possibilities that the lesson has for making the children conscious of the contributions of the past and of the opportunities for them to make contributions in the future?

The second thing we are going to look for and need in a democracy is self-direction, — the ability to initiate activities and carry them forward to success under one's own direction. It is very easy to test yourself on this. I can supply you with a test right away so you can tell whether or not you are developing self-direction in your schoolroom. Who sets the lesson? In a democracy, each person forms his own purposes and directs his own affairs. Why should it not be so in the classroom? Every school should give the children the opportunity to learn how to purpose.

As I see it, one of the big changes that is coming in education is a recognition of the need for a shift of emphasis. Formerly people in this country had initiative due to their frontier life. They went to school to get knowledge, for once they had knowledge, they had also the power to apply it. But our social conditions have changed. Life no longer supplies training in the exercise of initiative. The schools must now develop in the children power to apply knowledge as well as give them essential knowledge. From the kindergarten on thru college, the center of emphasis must be teaching children to purpose. The curriculum must be organized to give individual control. The thing we must do is to teach the child that his business is to grow in the ability to purpose. Purposing means self-direction.

We need to go a step farther. A person must have a critical faculty as well as ability to purpose. He must learn to appraise his own achievements. Who marks the papers in the class, your children or yourself? Every time a student in any grade of work from the kindergarten to the college has to submit his results to somebody else to know whether his work is satisfactory or not, we have autocracy instead of democracy — we have the opposite of self-appraisal. A great many of us are trying to develop this ability in the children so that they will be able to pass upon their own work. First, we are going to get the child to purpose to do certain things and then we are going to help him grow in his ability to say whether or not he has achieved his purposes satisfactorily. It is only as we change the achievement of the individual from unsatisfactory to satisfactory in school that we are going to get him to see the important things of life in the right relation.

Is not self-control also necessary in life? Are we always going to need to be under the control of policemen, or will we some day ourselves obey the laws which we ourselves make? The legislators we elect are our legislators. We send men to make the laws for us; they are our laws. But most of us do not act on that basis. Altho we teach the children differently, we do not follow out that plan in life. The opportunity for self-control comes wherever a group of people live and work together. Who controls in the schoolroom, you or the students? If the teacher controls, can there be much chance of development of self-control? The method used to control discipline is one of the best tests of the type of teaching going on in that schoolroom.

Finally, there is the vital question of coöperation. Our civilization has grown to the place where the life of a great city rests upon the division of labor. None of us live the complete life that the frontiersman used to live. We do not raise our own food, make our own clothes, build our own houses, etc. It is essential that we learn to coöperate. Do we realize that somewhere in this town this instant there is a man caring for the running of an engine, in order that we may have these lights? Do we realize the extent to which our activities and our lives are dependent on others? It is only when a coal or railroad strike ties up the business of the nation that we appreciate what cooperation means. Some of us are wondering at present as we read news from England, Italy, and Russia whether our civilization is going to hold out long enough to train a new generation that will be able to solve the tremendous problems that face us at the present time. The only way the child will ever learn to handle the machinery of constitutional government is in the schoolroom. It is the thing which we must give him in school above everything else, — the ability to work with his fellow-men in achieving common social purposes. So whether you are teaching history, reading, or writing, if I come into your classroom the things by which I ought to mark you as a teacher are determined by the extent to which you are giving those children the ideal of service, — developing self-direction, self-appraisal, and self-control, and are giving them the ability to use the machinery of coöperation.

Here are five types of control which are essential; and these are the things we must measure. Our measurements drive us into methods of teaching which are quite different from those in common use. It is easy to set up ideals of service, self-direction, self-appraisal, self-control, and coöperation, but the real problem, after all, is not to set up the ideal, but to put it across in the school system. What is the biggest problem to be solved? Is it not the training of teachers for service? If you had 4,000 teachers on your hands, all of whom have been trained under a different set of ideas, would you not think teacher training a mighty important problem? Those of us in this work believe it is one of the most important problems.

In our efforts at teacher training we have enlisted one of the latest tools that man has invented — the motion picture. The ideal of service and coöperation is not very much in evidence in the picture I am going to show you tonight. It presents a lesson in handwriting, so organized by the aid of measurement that the training in writing develops also self-direction, self-appraisal, and self-control. The film is called Democracy in Education.[1] It was made a year or so ago at the time when we were in the midst of the war. It is an attempt to put into dramatic and striking form those ideas which I have been trying to present to you tonight. The first part of the film shows the philosophy of the new method, the second part shows the actual lesson in writing to prove it is possible to organize school work very highly and not take away from the children the opportunity of entering into it whole-heartedly. I hope you will see that the whole scheme as it is worked out here gives teachers a chance to use measurements as a tool, as a means of furthering their teaching, and not to be controlled by it.

---

[1] The film entitled Democracy in Education may be obtained for a limited showing from the World Book Co., Yonkers, N. Y., for cost of transportation.

The Rating Scale System in Detroit

S. A. COURTIS

For many years we have had a rating scale system in Detroit. We feel that the one we now have is better than all the rest. We have experimented with different kinds of rating scales and we have found that before you can have a definite system it is necessary to have some form of reliable rating.

The type of scale which is in use in many places is the one on which are listed all the desirable qualities a teacher should have. The teacher is marked on her ability to supply subject-matter. We have worked these thru in Detroit, and our correlations do not work.

The time came when we decided that the mark of a teacher could not be decided in this manner. The teacher's rating is a subjective judgment on the part of the principal, an expression of opinion. The first kind of rating is a rank-order rating; the second, a general merit rating; the third, a description of that teacher — a grading in terms of certain outstanding qualities. This is more of a measurement scale. There is no correlation between the ability of the teacher and her grade on a test. We have some of our very best teachers who made low scores in tests under certain conditions — some of our poorest teachers who made high scores under certain conditions. We do not attempt to use educational measurements as a means of rating the teacher. The final opinion of the principal is better than anything else.

The principal has a rank order card on which are the names of the teachers in the order of their merit. For example, Miss K. Jones is the best of the teachers in a certain building, meaning the best in the judgment of the principal. So far as he can distinguish between Miss Jones and Miss Smith, Miss Jones is the better. We can thus show immediately where the principal of that building rates his teacher.

The next thing we do is to have the principal rate a given quality. All satisfactory teachers are put together, and all unsatisfactory teachers marked "D". Of three teachers, the extraordinary teacher is marked "A"; the fair, outstanding teacher is the satisfactory "C" teacher. The "B" teacher means the good teacher, between these two classes.

When you begin to grade your teachers it is very easy to grade A and C teachers, but when in doubt place her in the B group. All the principal has to do is to hold in mind these three grades.

On the other extreme we have the D, D+, and D— groups, in which go the teachers the principal gives up as absolutely hopeless. D means that possibly the teacher may do better another time. D+ is the case of the teacher who is unfit, but for whom the principal knows that nothing much can be done. The teacher who is old in service but who is being left behind, whom it would, however, be wrong to dismiss, is in this group.

When this system was put into work, requests came in from the principals for another mark. C— is the dumping-ground for many teachers — the teacher who is about to be married; the teacher who has been sick, but the teacher who will come back and be a good strong teacher. There are a number of those cases where this dumping-ground is needed.

Going back to the rank-order card, we can fill in the principal's estimate. The principal's judgment varies. Part of the work of our city is to bring the teachers to right standards of judgment. The minute we locate three teachers we can move the scale up and down the line if the principal's judgment is merely on the wrong level.

The rating and the rank-order are two forms of rating which check on each other. They are two different types of form. We find very often that rating in respect to each other is better than the final standard. But we need something more than this.

Here is an A teacher; here is another. How does the one resemble the first? If we need a principal, which shall we take? The ordinary scale of passing on teachers is a very difficult one to handle. It becomes very difficult to pass on these various elements of a teacher's ability. On the other hand, there are things we need to know to handle the question.

The following elements are considered essential to teaching ability: (1) vitality, (2) personality, (3) general intelligence, (4) social intelligence; dealing with professional work— (5) professional spirit, (6) professional ability, (7) professional leadership, (8) executive ability, (9) adaptability.

On this kind of card we have the usual division of 10 points. The rating of these teachers is in a summary of the elements named in the foregoing slide.

For the actual rating of teachers the army rating scale method is used. The instructions are to rate these divisions thru for each quality. When you rate that division thru, rate the one thru for excellent and very poor. You will find that the pictures of various individuals will rise in your imagination. Why? Because they have been made up from the statements of different principals until they strike upon someone you know. No principal can get thru that experience without coming in contact with some who are in excellent health, and some who are very low in vitality. Write the name of that person in and use that for a guide. For you and for me that name may not be the same. But both represent the best type of that division. Then, after writing in the name of the very best, write in the name of the very poorest. We have two elements of the scale. Without very much trouble, we can find a third person who will give us an average. Principals tend to come together on these averages. This scale tends to give us a fairly definite scale. If you will substitute from your own experience the names of definite persons, when you get thru you will have a complete scale which represents these five divisions.

The principal retains these scales. He keeps them from year to year and changes them only as he finds persons more suitable.

To apply this to a specific case, we have here a B6 teacher in Room C. There are 12 in the building. She stands third from the top, so she is third out of 12. The principal gives her a general rating of B.

The question now presents itself, How do you summarize these to get the rating? We don't. There is no connection between the rating and the description on the card. One is a good teacher because of her personality,— another good because of her adaptability. The value to be assigned cannot be combined by any summary.

In marking pupils, when the child is near the passing mark, we are likely to be liberal with the child. The principal does the same thing with the teacher.

The question now comes up, What are the factors which determine the principal's judgment? The thing that stands out more than anything else is for the teacher to contribute to the life of the school. The teacher who could attend with ability to the children in her own room and at the same time contribute to the other people in the building attracts the principal's attention more than any other type of teacher. This, I think, is a tribute to the type of work which is going on in Detroit.

This chart shows the correlation between the judgment of the principals concerning teachers and the elements prescribed as essential for a teacher:

Professional leadership	.87
Professional ability	.84
Executive ability	.77
Social intelligence	.69
Personality	.68
Professional spirit	.63
Adaptability	.63
General intelligence	.54
Vitality	.49

There is a deadline in the intelligence scale of teachers below which it is not safe to hire individuals. A result of not more than 125 on the Otis test renders the applicant doubtful; 80 means an almost sure failure.

If the people who are incompetent can be picked out from the beginning, it is our duty to the profession to see that only those who are competent to go on are in the profession.

The handling of such a scale is difficult. The scale is not satisfactory, but we have the means of making it so.

At the meeting of the Board of Education in Detroit last Thursday, the superintendent was authorized to issue two kinds of contracts. To the teachers whose work was satisfactory, a continuing contract is issued. To those whose work was unsatisfactory, a provisional contract is issued. The provisional contract can be used for only two years.

The suggestion has been made that there be some limit to the amount of salary given for service. Our minimum for elementary school teachers now is $1,500, for high school teachers $1,700; the maximums are $2,000 and $2,600, respectively.

If our scales are accepted as a just rating scale, the time will come when the next step will be payment for merit, so that the teacher who is rated A will receive an additional salary.

Any satisfactory scheme, to be workable, must be above the criticism of favoritism. I am quite sure that within one or two years we shall be able to work out a system, which, coupled with the opinion scale, will rate the teacher correctly and disinterestedly.

Analysis of School Subjects (Abstract)

Charles H. Judd

This is a plea for a type of study of educational problems which is more arduous than the making of tests or the comparing of scores. It is not very difficult to give 100 children a test in arithmetic or reading and to show from the results that some are good in each of these subjects, some are bad, and some mixed in their attainments. It is a very much more difficult task to follow up each case and find out what the score means. Until, however, we can ascertain this meaning of the results of tests, it is going to be difficult to interest teachers in the use of tests for improvement of teaching. If, on the other hand, we can guide teachers in the diagnosis of individual cases, tests will immediately take on a new significance.

As a first example of analysis, let us consider a case worked out by Professor W. S. Gray. He found a boy who was failing in fourth grade reading. The cause was the boy's failure to begin at the beginning of a line and systematically follow the words and phrases to the end. He rambled inside of the line.

A like example is to be found in the unsatisfactory method which some adults use in adding a long column of figures. They ramble around looking for easy combinations and lose some of the digits or repeat.

In the high school subjects there are examples. Pupils fail in geometry because they try to hold the propositions by memorizing, while in one of the other subjects, such as history, they learn by a better method and become interested in explanations.

In Latin tests it is shown that the third year course is very defective. The reason for this is to be found partly in the inappropriate subject-matter presented in the third year, partly in the vocabulary method of translation acquired by pupils in the earlier years.

Defective methods are in some cases due to the continuation in the later stages of a subject of a method of work which was all right at first, but is outgrown and defective for later stages. Creeping is all right in infancy, but not in mature life.

Again, analysis will often make us aware of the fact that training is dealing with mental powers which have nothing to do directly with the subject-matter of a course. Adding a long column of figures requires a power of sustained attention as well as knowledge of the addition tables. A lesson in history demands training in fluent reading. If the reading is slow and defective, it will interfere with the learning of the historical facts.

These examples show that we must go behind the results of tests. The tests will show children to be efficient or defective. Analysis is necessary to show the reasons which lie back of these test results.

Periods of Educational Development (Abstract)

CHARLES H. JUDD

ANY analysis of educational processes is sure to concentrate attention on the fact that children have very different types of capacities at different ages. The older child differs from the younger, not only in the amount of information which he possesses, but also in his modes of nervous and mental action. A mature nervous system is a network of highly organized tracts and a new impression coming into this organized system will be transmitted in a wholly different fashion from the same impression entering into the diffuse and little-organized nervous system of the younger child.

The general principal of periodicity was brought into educational literature in a very striking way by Hall thru his studies of adolescence. There are fallacies in Hall's treatment of the matter; perhaps the most striking is his contention that adolescence is a sudden and surprising something wholly new and unrelated to the child's immediately antecedent development. The fact is that adolescence can be understood only when we trace the full history of the child's development.

The best explanation of adolescence and the best basis for a full understanding of the process thru which a child matures can be gained by the study of the changes which take place in the process of digestion. The infant organism cannot perform the elaborate secretive processes which are necessary to digest solid foods and foods of great variety. The glands of the digestive tract constitute an elaborate battery of chemical stations and these glands can be brought under control only thru gradual use and adaptation to each other. While this process of coördination is going on the infant can digest only simple, predigested foods. The appearance of teeth and the slow growth of power to master the chemical complexities of digestion mark the arrival of the child at the advancing stages of maturity. Adolescence with all its expressions of maturity is the outcome of a long series of earlier adjustments on the physical side. The adult organism requires strong food, and the organism cannot make its last adult growth until the problems of digestion have been solved.

The analogy of digestion carries us over into the field of psychology. Here, too, there is a series of facts which show slow development. The first grade child cannot include in a single act of attention very many elements. The first grade child has no discriminating view of himself as distinguished from the rest of the world. The first grade child does not have the power to critically judge what is true and false. He must develop gradually thru many trials in the course of which he tries to adjust himself to other people and to the physical world about him.

The outline of a scheme of periodicity which will be helpful in constructing the school curriculum is as follows: Primary pupils are imitative, they have not established a high degree of self-assertiveness, they are readily led by the teacher and will readily acquire the fundamental social arts, especially reading.

The pupils of the fourth, fifth, and sixth grades are highly self-assertive; they are not especially interested in social example. They require a great variety of objective information to satisfy their personal interests.

The junior high school period is a period of social interest, wholly different from the primary period. The pupil is no longer imitative for he has passed thru the fourth, fifth, and sixth grades, where he has cultivated a high degree of self-consciousness. The junior high school period is one of self-conscious desire to know society. The social sciences are then appropriate subjects for the training of pupils of this period.

The senior high school period is one of partial specialization.

Without attempting to defend the foregoing as the only possible description of the maturing process, it may safely be asserted that some such outline must be arrived at as the basis of the course of study. The evolution of American schools has brought us to the point where we are going to give each pupil ample training. The selection of subjects for the schools is no longer limited to the few rudimentary subjects. We are now in a position to organize an elaborate course if we can discover the right principles of organization. The principle of organization which is here advocated as highly important is the principle of periodicity.

An Efficient Method for Measuring the Results of Instruction in Colleges and Schools

WILLIAM F. BOOK

I. An Efficient Method for Measuring School Accomplishment Defined

Two methods may be used for measuring the school accomplishment of a particular pupil or class, and the success achieved by a teacher or school: (1) By the use of standards set by the world and determined by a study of the school accomplishment of a large group of unselected individuals. This gives a sort of normal standard of achievement with which to compare the accomplishment of a particular individual or group. By this method and by the use of such world standards the school accomplishment of pupils and the success of teachers, classes, and schools have, as a rule, been measured in total disregard of the native mental endowment of the pupils whose scholastic achievement is being evaluated. The ability of the pupils to learn, or achieve scholastic success, has, as a rule, not been taken into account in our educational measurements. (2) A second and more pedagogical method for measuring the school accomplishment of a given pupil or class, and therefore the success of a teacher or school, is by the use of standards and norms which take into account the native mental endowment of the pupils with whom the teacher or superintendent must deal, measuring achievement in school and the success of teaching in terms of what may reasonably be expected in the way of scholastic accomplishment from pupils who possess their particular type and grade of ability. If we wish to evaluate the progress in school that can or should be made by a particular pupil or class, their native mental endowment must in all justice and for the sake of accuracy be taken into account. In no other way can we determine what ought to be achieved or demanded in school from such an individual or group. This is the true pedagogical method and should be used by all teachers in dealing with individual students or classes; and by superintendents in evaluating the work of a particular teacher or school or when comparing the results obtained in his own system with the accomplishment of another system in a neighboring city or state. For the latter may contain pupils possessing a very different grade of native mental endowment from that possessed by the pupils with whom he must deal and work.

In deciding what can or ought to be done for a particular individual or class or for individuals who vary so tremendously in native mental endowment as do the pupils of any class, community, or school, we need a method for measuring their accomplishment in school which will take into account the native mental endowment of each pupil or group, instead of a standard which is only applicable to a theoretically average individual or which is demanded by the world. Such world standards and the methods of measurement based upon them are still helpful for purposes of educational and vocational guidance, but they are wholly useless as a means of showing what may reasonably be expected in the way of school accomplishment from a particular individual or group. They show what degree of accomplishment must be achieved by an individual if he is to be successful in a given line of work or if he is to make normal or unusually rapid progress. They are, therefore, valuable for certain purposes, but such a method can never be used as a just or accurate method for measuring the *school accomplishment* of pupils or to *evaluate the success of a superintendent or teacher* because it fails to measure what is being achieved in the light of what actually *can* or *ought to be* accomplished by individuals or groups who vary so tremendously in ability to learn. Such a true pedagogical method for evaluating school achievement is a necessary instrument for any teacher or superintendent who desires to conserve and cultivate to the fullest extent the capacities and talents of all his pupils.

II. Need for a More Just and Efficient Method of Measuring Accomplishment in School

Some of the reasons why such a pedagogical method should be used to measure the progress of pupils in school and the success achieved by a teacher or superintendent should perhaps be pointed out in detail. First and foremost of these is the fact of the marked individual differences in ability to learn and do which we find when we compare individuals belonging to any homogeneous group or to similar groups in different communities and schools. The native mental endowment of pupils belonging to different classes in the same school system or to schools located in the same county or township varies tremendously. Three years ago we made a mental survey of all the schools in one Indiana county and found a rural school in one township where 19 pupils were present on the day the tests were given; 11 of this number, or 56 per cent, were feeble-minded. Twenty per cent of the school population of the entire township were found by individual examination to be feeble-minded. In another township 11 per cent of all pupils present in school on the day the tests were given were feeble-minded. In a third township less than 2 per cent of the total school population was feeble-minded. It should be added that this is perhaps a very unusual case, but it makes clear the fact that the school accomplishment in township 7 where less than 2 per cent were feeble-minded cannot in justice be compared with the results obtained by teachers in township 1 where 1 pupil out of every 5 was feeble-minded.[1] We need a standard for measuring school accomplishment which will enable us to determine what may reasonably be expected *from* and what should be done *for* pupils who differ so widely in native mental ability.

In our state-wide mental survey of high school seniors we found differences in the native mental endowment of the senior classes representing individual high schools which were equally marked, even when these schools were of the same size, and located in the same county, township, or city.[2] This fact is strikingly shown in Figure 1 which represents the senior classes

---

[1] Compare study by W. F. Book, "Variations in Mental Ability and its Distribution among the School Population of an Indiana County," Proceedings of Fifth Annual Conference on Educational Measurements, Indiana University, Bulletin of the Extension Division, Vol. IV, No. 4.

[2] Compare study by W. F. Book on "The Intelligence of High School Seniors" published by the Macmillan Co., November, 1921. 350 pp. See especially Chapter XII. Their courtesy for the use of these figures is gratefully acknowledged.

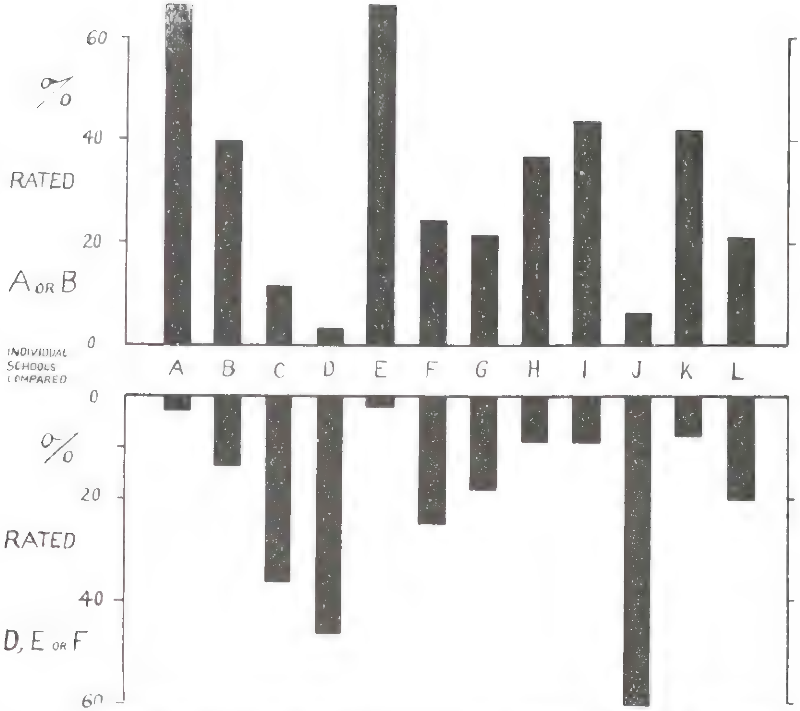

FIG. 1. Percentage of seniors from all high schools in state ranked 3 who possess the highest (A or B) and the lowest (D, E, or F) grades of intelligence found among high school seniors.

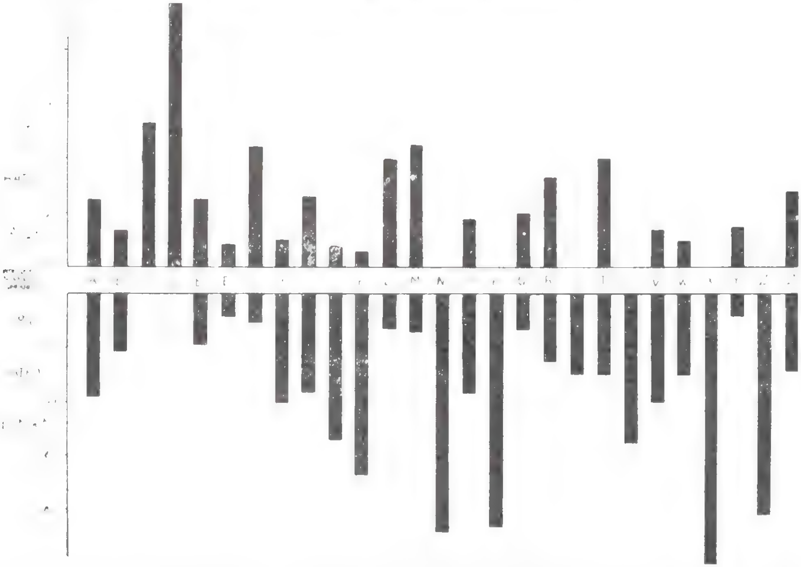

FIG. 2. Per cent of seniors representing each of the first 26 high schools in the state ranked 6 (alphabetically arranged) who made an intelligence rating of A or B and D, E, or F.

for 1918-19 in all the high schools of the state ranked 3. It shows the percentage of pupils belonging to each senior class who made the highest (A or B) and the lowest (D, E, or F) intelligence rankings found among high school seniors in the entire state, and, as may readily be seen, the differences in mental endowment of these senior classes are strikingly great.

Figure 2 compares the native mental endowment of the senior classes in the first 26 schools of rank 6. The 90 schools in the state belonging to this rank were arranged in alphabetical order and the first 26 schools taken to show the variations in intelligence found among their senior classes. The figure shows the per cent of seniors who made an A or B ranking on the mental tests and the percentage ranking D, E, or F. In some of these schools the entire senior class made scores which placed them above the median for the entire state while in other schools the entire senior class ranked below the state median. Differences equally as great as those shown in this chart would appear if the last or middle 26 schools belonging to this rank had been selected for the comparison, or if we had compared the intelligence ratings of the middle 50 per cent of seniors belonging to each of these schools. See Figures 3 and 4.

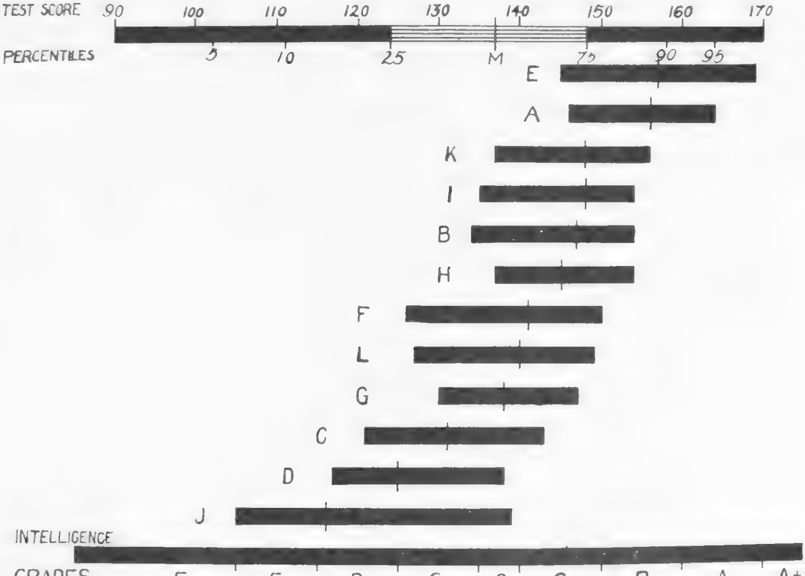

FIG. 3. Scores made by middle 50 per cent of the seniors in each high school ranked 3.

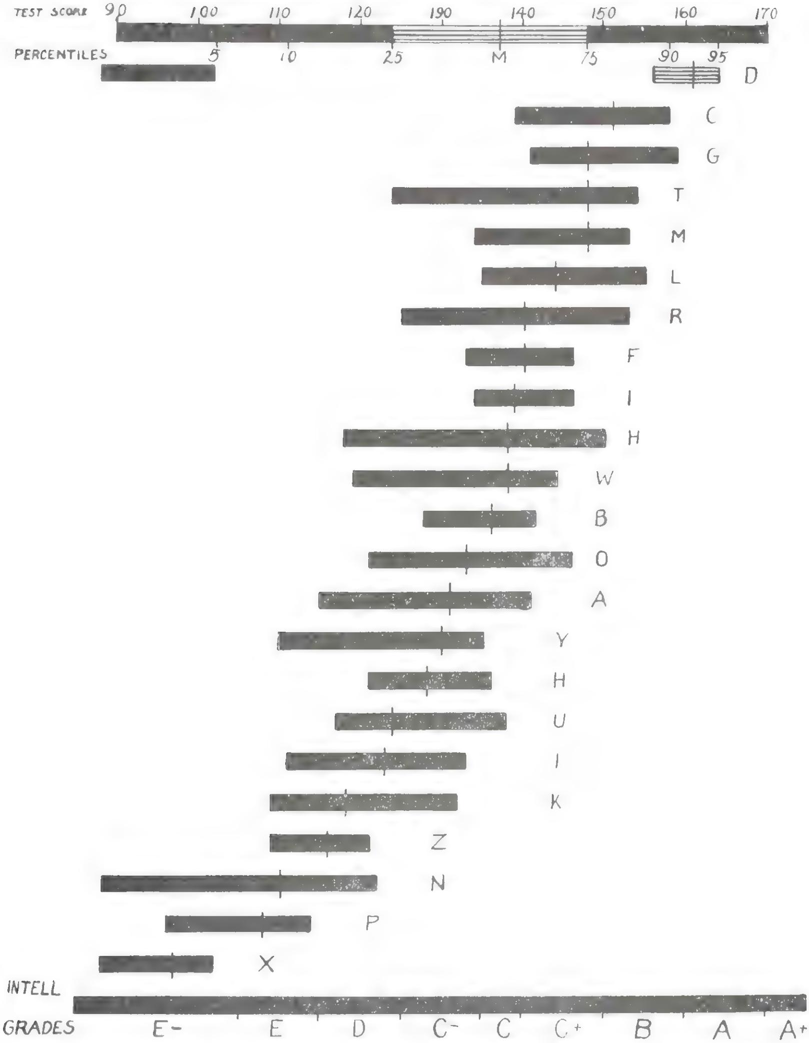

FIG. 4. Scores made by middle 50 per cent of the seniors belonging to the first 26 schools of rank 6. Figures at right of the bars are key letters and refer to individual schools.

Similar differences appear if we compare schools of the same size and rank located in the same county or city. Figure 5 compares all the high schools located in four Indiana cities A, B, C, and D. The rank or size of the schools is indicated by the figures placed on the vertical bars, representing the several individual schools. It will be seen from an examination of this figure that the senior classes in cities A and C show a marked difference in native mental endowment; the two schools in city D are about equal in mental strength; two of the three high schools in city B are about equal in native mental ability, while the third school ranks much higher in the percentage of seniors who belong to the highest intelligence ranks.

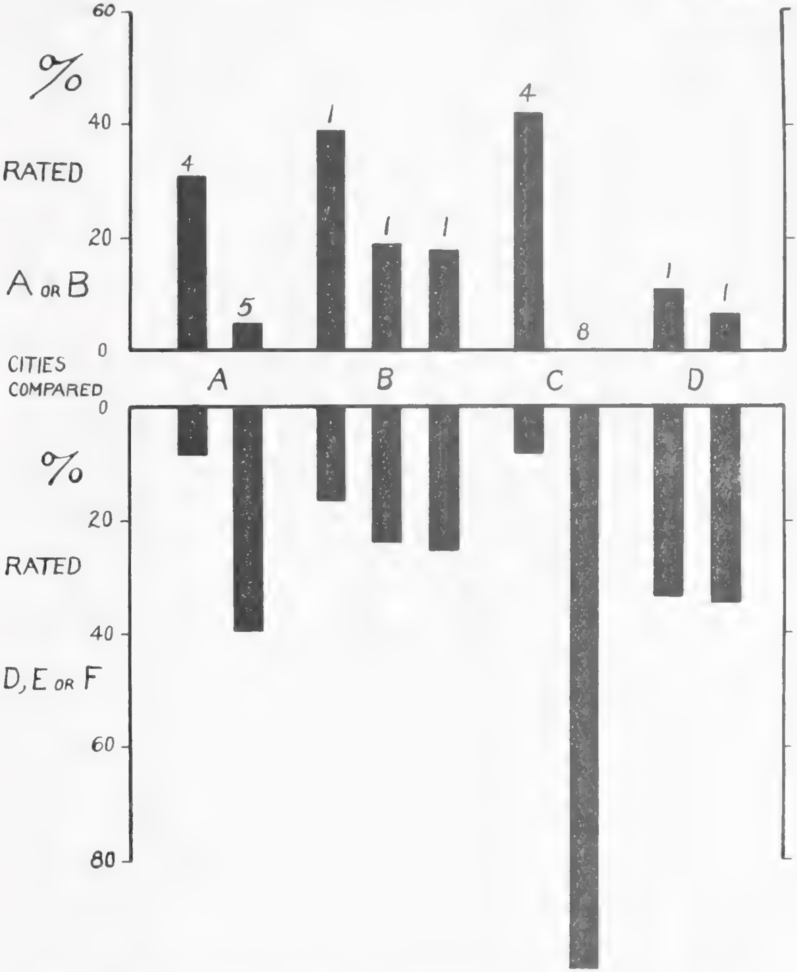

FIG. 5. Per cent of seniors in the high schools of four cities (A, B, C, and D) who made the highest (A or B) and lowest (D, E, or F) intelligence rating on the tests.

The same variability in native mental endowment is shown if we compare schools in the same county. Figure 6 compares the senior classes of all the high schools tested in three Indiana counties, selected at random from the 92 counties of the state. The figures placed above the vertical bars indicate the rank of the school. The length of the black bar above the horizontal line shows the percentage of seniors belonging to the school who made the highest mental rating obtained by any high school seniors in the state, while the distance below the horizontal bar indicates the proportion of seniors in the school who obtained the lowest (D, E, or F) mental rating made by high school seniors in the state. It will be seen that the schools in county F, if considered as a whole, are decidedly superior to the schools tested in county G. The same remarkable variations from school to school are shown if we compare the schools in any county.

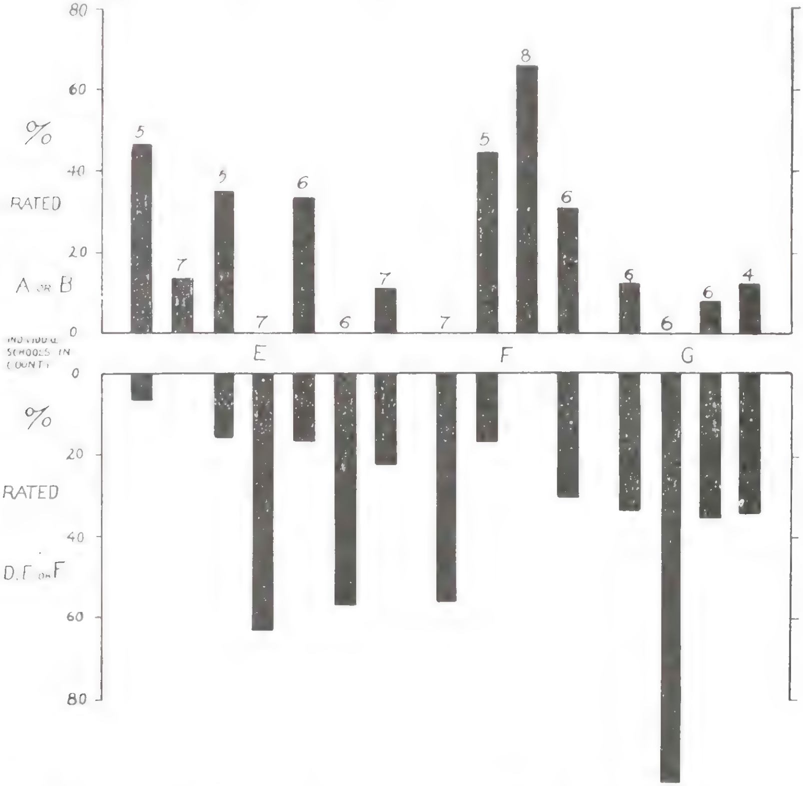

FIG. 6. Per cent of seniors representing individual high schools in three counties (E, F, and G) who made the highest A or B and lowest D, E, or F intelligence ratings on the tests.

It would clearly be unjust to both pupils and teachers to expect the same quality or amount of school work from senior classes which differ so widely in native mental endowment. In fact, it would be impossible to get the same scholastic results in these several schools.

A second reason why a more pedagogical method should be used to measure the accomplishment of children in school is that in no other way can we determine what may reasonably be expected *from* or done *for* pupils who vary so greatly in native mental endowment. A comparative study of the intelligence scores and school accomplishment of pupils in the public schools and colleges has shown that it is as a rule the most capable individuals who are really retarded and therefore most poorly served by the school because they are encouraged by our present methods of evaluating their work to do less than their best. In other words, we are not getting returns from our brightest students commensurate with their ability. They are often made to feel by our present methods of evaluating school accomplishment that they are doing very well when they are doing only a little better than average work, when because of their exceptional ability they should be doing a *very superior* grade of work. Students who possess a more mediocre grade of ability are, on the other hand, often found to be advanced more rapidly than would be normal for them. And because they do a poor grade of work when compared with their superior classmates they often become discouraged, and paralyzed in their efforts because they are not told or shown that they are, in reality, doing a very commendable grade of work for them. As a result of such methods of evaluation many of the brightest students in our colleges and schools are really encouraged to form habits of work far below the level of their best achievement, while extra time is spent on the duller students who are already doing a fair or superior grade of work when we take into account their native mental endowment. There is great human waste because we try to bring the dull up to standards of achievement which are clearly impossible for them, while the more gifted students are allowed to form habits of mental laziness because they are not given the kind of guidance and help which their superior ability requires.

A few facts should be given in support of this statement. In our state-wide mental survey of high school seniors a careful study was made of the intelligence rankings of all high school seniors in the state who had been accelerated, normally promoted, or retarded in high school and at some time during their elementary school course. This study revealed the astounding fact that the brightest seniors in the state were never selected for special advancement either by the high school or in the elementary grades. Seniors with a high average or slightly superior grade of intelligence as high school seniors go were often advanced more rapidly than normal, but none of the seniors whose intelligence score placed them in the highest one percentile group for the state were found in any of the accelerated groups. These more gifted individuals either had not been located by the school, or if located, they had not been permitted to complete their course in less than normal time. Of the seniors who possessed an A+ or A grade of mental ability only 6 and 7 per cent respectively were advanced more rapidly than normal in high school, while 91 and 90 per cent of these most superior students were only normally advanced.[3]

Our study further showed that most high school seniors who possessed the most inferior grades of intelligence found among high school seniors (93 to 97 per cent) had been regularly promoted in high school, while less than 6 per cent of our total group finished their high school course in less than normal time. Our results therefore clearly show that the schools are not adapting themselves to the native mental abilities of their students as well as they should; that they are accelerating too few students; and that the school officials are unable to select the most intelligent students for such special advancement or to induce them to do the quality of work which is commensurate with their ability.

---

[3] Compare the writer's complete report, "Intelligence of High School Seniors," The Macmillan Co., Chapters V and VI.

Summarizing the results of our entire study on this point we found that more than twice as many students were kept on their high school course four years as possessed an average grade of intelligence for high school seniors; about seven times as many possessed a superior grade of intelligence as were actually permitted to shorten their course, while less than one-third of the group possessing very inferior grades of intelligence were actually retarded. We may conclude that many students in our schools are working far below the level of their best attainment and so acquiring habits of laziness and inefficiency because their school work is ill adapted to their mental strength and needs. The remedy would seem to lie in a method of evaluating school achievement in terms of mental capacity or ability to learn; to apply in all our educational work the principle set forth by the Great Teacher in His parable of the talents. In no other way can the capacities of our students be fully conserved and the work of the school made truly economical and efficient.

Similar facts have been revealed by our study of the intelligence ratings and scholastic records made by college students and children in the elementary grades of the public schools. Last October Mrs. Pressey, research fellow in psychology, made a special study of the school accomplishment in arithmetic of two fourth grade classes in the Bloomington schools. The actual progress made in the fundamental processes of arithmetic was compared with the pupils' actual ability to master this work. It was found that only a low degree of correlation existed between ability to do this work and their actual achievement in school; the coefficient of correlation was only .28. A careful analysis of the individual records showed that the lack of correlation was due chiefly to the fact that the brightest students had fallen, in their achievement, far below the point which their native mental ability indicated that they should attain. By carefully measuring the school achievement of each of these pupils and representing the results on a card in a manner which would picture to each pupil his actual scholastic accomplishment in terms of what should be achieved and showing him in which particular respects his record was below what it should be and at what points it was normal or high, and by presenting such records to these children at intervals of two weeks for a month and a half, the achievement of the brightest students in the class was sufficiently raised to increase the coefficient of correlation between school attainment and their mental test scores from .28 to .85. For the control group in the same grade and school no such results were obtained. The coefficient of correlation was increased only 10 points, from .35 to .45. Not only was the school record of the brightest students thus materially raised and these pupils shown what they really should accomplish, but the less fortunately endowed members of the class were for once in their life proud of their achievement. As one of them enthusiastically remarked: "This is about the first time in my life that I have been told I was doing well in school." But what is perhaps still more significant, the most superior members of the class were for the first time in their life forcefully reminded of the fact that *they* were not doing as well as they ought to do. These brighter children had been well satisfied with their record in school until they received the first report of the experiment. They had always stood at the head of their class and would have received no particular encouragement to do better work from the old-fashioned report.

A similar situation has been revealed by a study of the scholatic records made by 300 college students evaluated in terms of their intelligence ratings obtained by giving the Indiana University Intelligence Scale. Our results clearly show that it is as a rule the brightest students who are doing less than they are able to do. In evaluating the scholastic record of these students the principle set forth in the parable of the talents should as far as possible be applied in every individual case and their school accomplishment measured in terms of what they are really *able* to *do.* This is the only method that is accurate and just and the only way in which these students may be helped according to their needs.

III. How the Work of a Pupil, Teacher, or School may be Measured in this More Accurate and Pedagogical Way

For more than a year we have been making a study of the scholastic success attained in college by the brightest seniors who graduated from the high schools of the state in 1919, and have given some attention to devising a technique which would enable us to measure such scholastic success in terms of their native mental ability to do the work. It is obvious that we cannot measure the school accomplishment of students in terms of their ability unless we combine the results from a mental survey with the results obtained from the most reliable measurements we can make of the school achievement of the particular individuals or groups whose scholastic accomplishment we desire to evaluate. A careful mental survey must be made and the results studied in relation to the most accurate measure we can get of the school accomplishment of our group.

In devising such a pedagogical method of measurement it would be desirable if an accomplishment quotient or index of learning ability could be obtained which would indicate the relation which exists between a student's native mental endowment or ability to learn and his actual scholastic success. If such an achievement quotient or learning index were devised it could be used to determine what a student, with a given grade of ability, *could* and therefore *should be expected* to accomplish in school.

In the earlier stages of our mental measurement work it was thought that a reliable intelligence score really constituted such a measure of accomplishment. As a result, intelligence ratings have often been used as a guide in the promotion of pupils in school, as a test for admission to college, and in many universities as a basis for awarding scholarships and fellowships to the most capable students. But careful studies have shown that other factors besides intelligence enter into the making of a successful record in school and life; that quite as much depends upon the individual's attitude, his application and interest in the work, his interest in his own success and improvement as upon his intelligence or mere ability to do. These studies have made it very clear that we do not yet have any reliable method for determining what success a given student will attain in school or life and that the problem of obtaining such an achievement quotient is exceedingly difficult and complex. Nevertheless, it is important to evaluate as far as we can what a pupil does in college or school in terms of what he is really able to do and to devise the best standards and methods we can invent for measuring the results of our teaching in this true pedagogical way.

Guiding Principles. In attempting to devise such a standard and method for evaluating school accomplishment we have kept in mind the following principles and aims:

1. Students should be shown not only what they are actually accomplishing but what they are really able to do so far as this may be indicated by a mental test score supplemented by teachers' estimates of their ability. They should be given reports showing not merely what they have achieved but what they are able to achieve so that they may have a definite goal toward which to work, a goal which they could actually attain if willing to work.

2. The report should be sufficiently diagnostic in character to show each student in what particular subject he is *below*, *at*, or *above* the standard of achievement which is normal for him.

3. The method must be accurate, thoroly practical, and just to the pupil.

4. The report should be so worked out and presented that it would encourage the student and stimulate him to do better work.

5. The report should be presented in a way that the practical value of setting such a definitely attainable goal before the student could be measured and the degree to which it stimulated him to do better work accurately determined.

6. Comparison should be made with his own class, or group, rather than with standards which are more fixed in character or universal in application.

Method Described. With these principles in mind a new form of report card was devised which it was thought might accomplish most or all of these purposes. Seven University classes, 300 students, were given the Indiana University Intelligence Tests and the range in mental rating made by each class determined. The median and 5 percentile scores for each class were next determined and the results represented on 5 x 8 cards in such a manner that the median score for the class and the individual score made on the test by each student was shown in its appropriate relation to the ratings made by the entire class. To accomplish this, the card was divided vertically into 20 equal parts, each part representing the range in mental test score made by each 5 percentile group, ranging from the lowest or 5 percentile group to the highest or 95 percentile group. Each individual's intelligence score was then represented on the student's individual report card in its appropriate place in this total range of mental ratings.

The card was divided vertically on its right hand side into 12 equal divisions, representing the various scholarship ratings given to students in their college subjects at the end of the semester. The card was next divided horizontally into 7 equal divisions, and the scholastic achievement attained in each subject as indicated by the professor's term marks represented on these vertical lines at the points indicating the scholastic rating actually attained in those subjects. The scholastic standing in each subject was therefore shown for each student in relation to his own mental rating, and in relation to the median ability of the class. These points were then connected by a solid black line, and the average scholarship in all subjects determined by calculating on a credit hour basis the percentage which each student obtained in all subjects studied during the semester. This average scholarship rating was indicated by a dotted horizontal line on the card. Such individual report cards were prepared and given to the instructors and to each member of the class. They showed the median score for the class' the individual's own mental rating, his average scholarship rating, and his actual accomplishment in each subject studied during the semester as indicated by the marks given him by his professors at the close of the term. See sample report card in Figures 7 to 10 below.

These individual report cards were then presented to the students with the request that they be carefully studied and that they explain as far as possible any discrepancy which might appear in their individual records between their college accomplishment as indicated by their term marks, and their relative standing on the mental tests. Two groups of college students were studied: 137 juniors and seniors in the experimental psychology classes, and 127 sophomore students taking the course in elementary psychology.

A careful examination of all these report cards showed that a great many students made a scholarship record which was much below the achievement which their mental test score indicated they should attain. (See Figure 7.) This was particularly true for the sophomore students who are taking the required work, and for the brightest students in the several classes. Some students did better in their college work than one might have expected from their mental test scores. That is to say their scholarship line was above the line indicating their relative standing on the mental test. (See Figure 8.) Other students made scholarship records which corresponded very closely to the achievement which one might expect from the relative standing on the mental test. (See Figures 9 and 10.

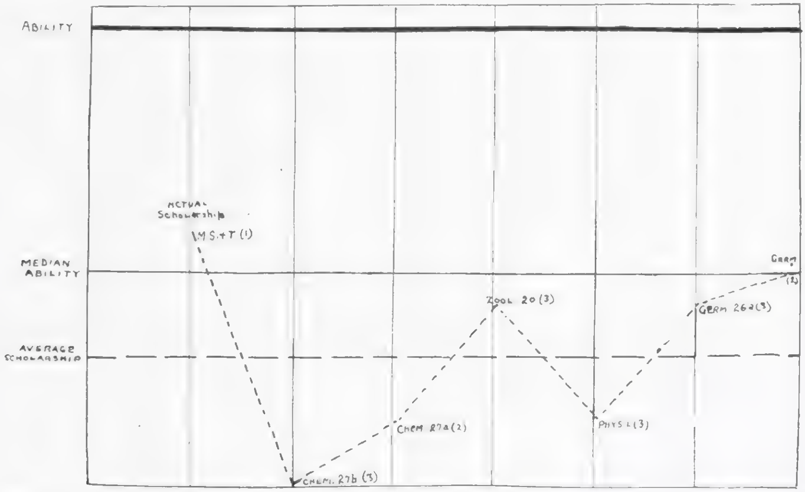

FIG. 7. Record of very superior student whose scholastic rating is very low.

In order to determine the cause or causes for these discrepancies from the students themselves and to ascertain the effect which such a report might have upon their future work, an added experiment was devised which aimed to determine these facts. The students were asked to explain as specifically as they could the cause or causes of any discrepancy which might be shown in their own case between their mental ability as indicated by their rating on the mental test and the actual accomplishment in their college courses. They were also asked to state in what specific ways, if any, such a report of their college work would benefit them; and told that their scholarship records for the present semester would be placed on these same cards and studied in relation to their previous scholastic records and intelligence score. In this manner we hoped to get some indication of the actual effect of such reports upon a student's college work.

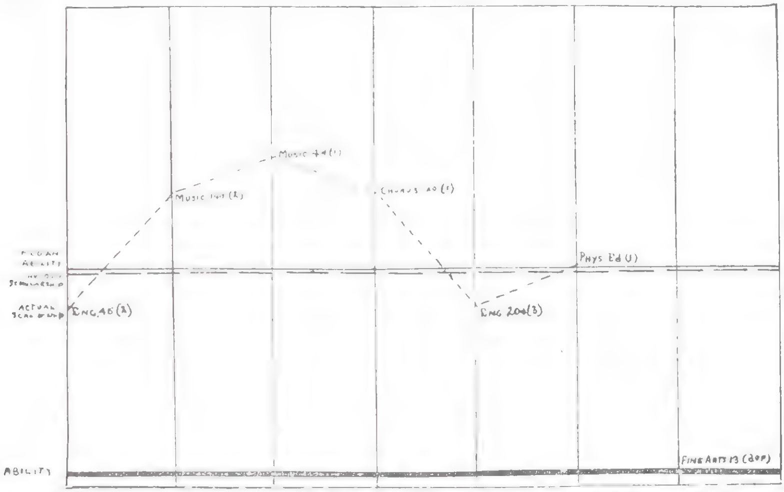

FIG. 8. Record of student of low intelligence rating who nevertheless is doing average college work.

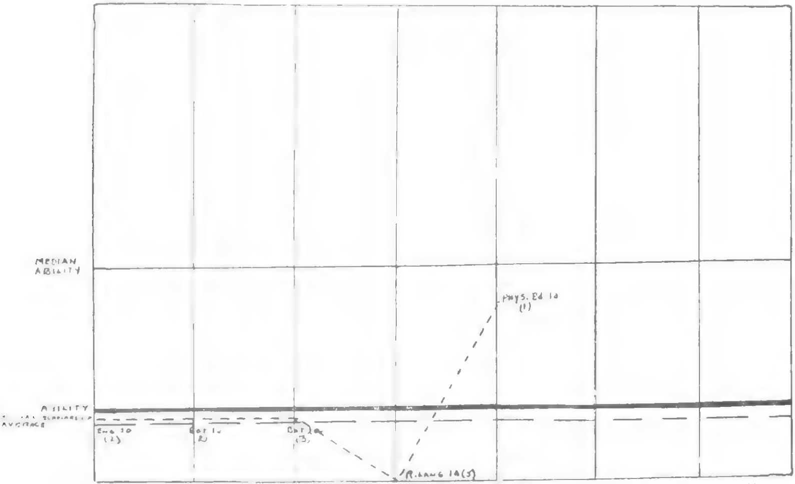

FIG. 9. Record of student wiht low grade of mental ability for college students whose scholarship standing coincides with his mental ability.

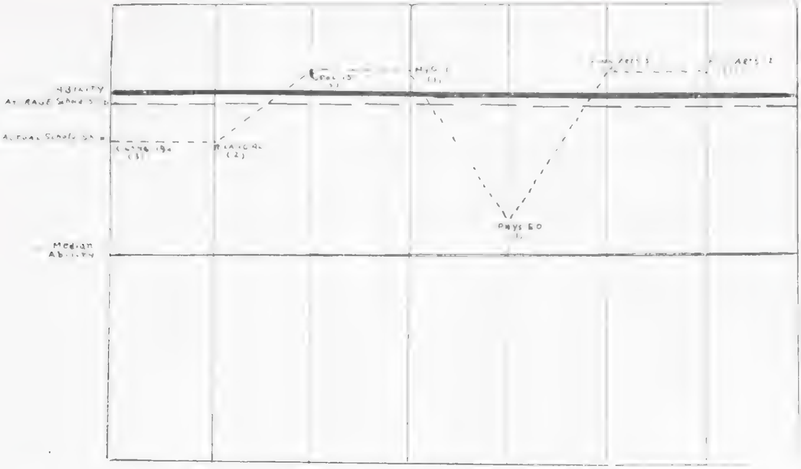

FIG. 10. Record of superior student who is doing a superior grade of work in college.

IV. Conclusions

Our conclusions as to the value or effect of applying this pedagogical standard of measurement may be briefly summarized somewhat as follows:

(1) Many of the brightest students will be informed probably for the first time in their school career that their achievement is not up to their highest standard of accomplishment and so will be stimulated and encouraged to do better work. One student said the record showed that he should do very much better work, but that he had been told this before. To many students, however, it comes as a new and striking revelation. Several stated that they had never been shown in a definite way what they really should accomplish in school.

(2) It will stimulate all those whose school accomplishment is below their ability index. This statement was made by about 90 per cent of the students and has been borne out by work done in the elementary schools.

(3) Pupils possessing only average or mediocre grades of ability will also be encouraged and stimulated, since many of them will for the first time in their lives be shown that they are in reality doing excellent work even tho their achievement is not up to that of the most gifted persons in their class.

(4) The method will enable us to study the causes of success and failure in school from a new angle, and will help to determine the characteristics other than intelligence which enter into the making of a successful record in school. When all these characteristics have been determined and their importance demonstrated they may perhaps be cultivated and assigned their true place in our scheme of education. We do not need diagnostic tests which will show what pupils *will* do in school, so much as we need tests which will show what they are really *capable of doing* so that we may know what in the way of school accomplishment may reasonably be demanded from all our students.

(5) Such a method will also enable us to measure the results of instruc-

tion in a more efficient and economical way. In evaluating the work of a teacher we must combine the results obtained from achievement tests with the results of a mental survey and learn to judge what is being accomplished by a particular individual or class in the light of what they are really able to do. It is not only unjust, but pedagogically indefensible to judge the results obtained by a teacher without regard to what can and therefore what ought to be achieved.

(6) Such a method of measurement will also give us a new technique for studying retardation. The best and perhaps the only way in which a superintendent can determine the extent to which his schools and teachers are succeeding in adapting the work to the native mental endowment of their pupils is by means of a mental survey. This may be done by determining the grade of intelligence which the pupils possess who have actually been accelerated or retarded by the schools. The intellectual ability of those accelerated may thus be compared with the native mental endowment of those who have been retarded or only normally promoted. If the brightest students are not making good grades on their school work the reasons for such failure should be immediately determined and the proper adjustment made. The method will also reveal whether the brightest students are advanced as rapidly as their mental ability warrants, and whether or not the dullest students are promoted along with the brightest or even ahead of them. In schools where the most careful mental surveys have been made we regularly find that many of the brightest students have not been advanced by the school as rapidly as their ability seems to warrant, while the dullest students have been advanced more rapidly than their mental ability indicates that they should be. In other words, it is the brightest pupils who are really retarded, because they are found to be from 1 to 2 years behind the grade to which their mental rating indicates they should belong.

(7) Combining the results of mental surveys and achievement tests is, therefore, one of the best ways to measure the efficiency of an entire school system. Mental surveys have revealed marked group differences between classes or sections in the same school system, and between schools located in different communities. The mental survey will enable a superintendent first of all to compare the intellectual capacity of different sections or classes in the same school grade and so to determine what in the way of school accomplishment may reasonably be expected. In this manner the work of a teacher may be properly evaluated, when without such intelligence scores all sections or classes would be judged by the same standard, which in the light of the group differences which have been shown to exist would be manifestly inaccurate and unjust. It is only by the use of mental tests that a superintendent may obtain a real factual basis for comparing the accomplishment of his own system with that of another system or to evaluate the results obtained by the teachers in different buildings. Intelligence tests, therefore, become not merely an instrument in the hands of a superintendent for making adjustments to class, school, and community differences, but a necessary means for making accurate pedagogical measurements of the results that are obtained in his schools.

(8) This method of evaluating school achievement will also, the writer believes, make our students more successful in life. It will show them while still in school more nearly what they ought to accomplish, and so will help

to develop in them habits of making the most of their abilities. They have often been told and urged to do this, but if such a means of reporting the success actually attained in their work were devised and used they would be shown by actual curves and charts to what extent they are really living up to the highest level of their ability. If this matter is left to take care of itself as is at present the case, many of the very brightest students in our schools will form habits of work far below the best level of their achievement and so hamper their chances for making the highest possible success in life. For it may be assumed that they will work in life about the way they have worked in school.

(9) Lastly, such a method of reporting accomplishment in school will help to conserve to the world the talents of the brightest students because they will be stimulated to make the most of their capacities and talents. This has already been shown to be the case in one school grade. Moreover, 90 per cent of the students whose college record was studied believe that such a method of reporting their school accomplishment would produce such an effect.

V. Summary

Summarizing in a word the main points of the paper, we have called attention to two kinds of measurements in evaluating school accomplishment: (1) The method which measures school achievement entirely on the basis of the results obtained, evaluating such results by standards set by the subject to be mastered or by the world. This method is valuable, among other things, as an aid for educational and vocational guidance because it uses the standards and achievement set by the world. (2) A more accurate and just method of measuring school achievement is to evaluate it in terms of the native mental endowment of the learner or in terms of what may reasonably be expected from individuals or groups possessing a given grade of native mental ability. The need for such a pedagogical method of measurement arises from the fact that we find such marked individual differences in native mental endowment among individuals and groups of individuals located in different communities and schools, and from the fact that the educational and life needs of individuals who differ so greatly in ability cannot be truly met unless we know what in the way of school achievement may reasonably be expected and therefore demanded from them.

As one possible method of evaluating school accomplishment on the basis of ability, we have proposed a new type of report card which will show a student's capacity to learn, as indicated by his intelligence rating, in relation to the median, best, and worst ability of a class or group, also his school marks and scholarship standing in relation to his ability to do the work. Such a report might be used merely for the information of the superintendent and teacher or it might be given to parents and pupils so that each pupil could see to what extent he was accomplishing what might reasonably be expected from him in the way of achievement in school. Such a report would reveal to the teacher what pupils were making the most of their capacities and those who were accomplishing less than they should. It would also enable a superintendent to evaluate more accurately and justly the work of his teachers and the efficiency of his schools. It would tend to encourage the student who is really doing his best but who cannot attain the results obtained by the most capable pupils in his class. It would also stimulate those who possess very superior ability, but who are doing only a good or average grade of work to do superior work as they should, and so develop habits of application and work commensurate with their ability, thereby conserving to themselves and to the world their superior endowments.

On the basis of the facts and data already available and notwithstanding the many weaknesses of mental tests and the inaccurate measures we now have for evaluating school accomplishment, we may nevertheless conclude that some such method of reporting the school accomplishment of pupils as is here outlined is worthy of being given a serious trial. It is thoroly practical and gives promise of yielding very important educational results.

In conclusion, I wish to say that the entire problem raised is being studied from various angles by the speaker and several advanced students in the department of psychology. We would be grateful for any criticisms of our methods and plans and for any suggestions about the probable value of this type of research. We should at any rate like to hear the problem and method fully discussed.

Measurement of Progress in English in the Upper Grades

SIDNEY L. PRESSEY

AT the conference last year, I presented briefly to you the first form of an examination for investigating fundamental achievement in the eighth grade. This examination, which we have called Scale of Attainment No. 2, consists of tests in reading vocabulary, the mechanics of English composition, American history, and arithmetical reasoning. You have doubltess all seen the blanks, which have been on display along with our other tests. It chanced to come about last fall, as the result of a conversation with a somewhat pessimistic-minded professor in another department, and as a result of observation of my own with regard to the shortcomings of the average college student, that we tried this scale for the eighth grade — a scale dealing with eighth grade subjects, in an eighth grade way — on several classes here at Indiana University. The results were, in some respects, distressing. It is surely unfortunate that a college student should know the answers to no more than 19 questions out of 40, dealing with facts in the history of his country considered sufficiently important to be put in the eighth grade curriculum. For when we are told that the War of 1812 was with Mexico, that the Louisiana Purchase was made by Monroe, that the adoption of the Missouri Compromise was brought about by Calhoun, that the political campaign of 1896 was concerned chiefly with reconstruction, or that the President of the Southern Confederacy was Andrew Jackson, one cannot but feel but that somewhere, somehow, sometime, there is, or was, something wrong. I could multiply to you examples of such historical absurdities, but surely to no great profit. In fact, I believe there is an ancient pedagogical law to the effect that a repetition of error may lead to its final acceptance. And I should be most unhappy if I should insinuate into your minds the idea that the chief officer of the judicial branch of the government is the President, that the constitutionality of acts of Congress is decided by the Senate, or that the Lewis and Clark Expedition explored the Great Lakes. I must not leave you with the impression, however, that all the errors occurred in history. The particular class to which I now have reference (made up largely of upper classmen) informed me that a gleaner works in an asylum, that a mosque is a kind of an amusement, that a blunderbuss is a kind of animal, that a curfew is rung at midnight (perhaps a vision of things hoped for), that a sylvan place is found in oceans, that a grenadier is a diplomat, and that a pathetic story is very profane. These words, mind you, are all taken from the glossaries of eighth grade readers — and are forgotten, apparently, soon after the student leaves the eighth grade.

The mistakes in arithmetic and in English grammar are less dramatic, but no less painful. It must be remembered that the college students, beside having had the benefit of the four years high school and the one year or more of college training, are also the selected pick, presumably, of the eighth grade group. Nevertheless, there are cases in the last three years of college who go below the eighth grade median, in eighth grade work.

Perhaps you may feel that specific words of English vocabulary, specific types of arithmetical problem, or specific facts in the history of the United States are not supposed to be remembered; they are supposed only to do their part in the educative process and then to disappear, leaving only some vague general benefit behind them. One then wonders if this vague benefit might not be procured with material *worth* remembering. But, however that may be, there is at least little room for argument with regard to the assertion that ability to write correct English is of continued practical importance. Work in English is also the one type of work which is carried fairly consistently from grammar school, thru high school, into the University. So we should expect a gain, from eighth grade to college, in the fundamentals of English composition. The English Department here has been much interested in this matter, and has very kindly given me the opportunity to do extensive work in the freshman English classes. The general discouraged attitude of the English instructors, with regard to preparation in the fundamentals of good usage, was distinctly justified. In fact, there appeared to be less gain, from the eighth grade to college, in ability to use English correctly than in any other subject investigated. And in both eighth grade and college the most elementary and inexcusable blunders appeared in English.

It seems to me very decidedly that the responsibility for this most unfortunate situation rests primarily upon the grammar school, first, and then the high school. It is the feeling of college teachers in English (and a feeling with which I very heartily agree) that, by the time a man or woman reaches college, drill in spelling, grammar, and punctuation should have resulted in established habits of correct usage in these respects. To this opinion may be added another, that much as few people develop their handwriting during their college course (the habits concerned being already firmly established and hardly conquerable at this age) so, few people change essentially or permanently their grammatical habits, or their habits of punctuation or spelling, from the time they enter college or earlier. If drill is to be effective, it must be during childhood, and in the public schools. English teaching has, apparently, lacked definiteness of aim. There has been lacking intensive and sufficiently specific drill. There has been an unfortunate diversion of attention to literary values largely beyond the comprehension of the average child. As a result, the schools are constantly turning out children who progress into college, or out into the wide, wide world, with an astounding inability to express themselves accurately or correctly in written work and, often, with only a cheap smattering of sentimentalism, or an early adolescent disgust of all things "literary", as the only lasting end product of the "inspirational part" of the teaching.

It was primarily from the test results I have just mentioned, and from the results as thus interpreted, that the present study grew. When such a situation is suspected, the first need is for a definition of that situation by means of some exact measurement. Such measurement would also seem the natural point of departure in any effort at improvement. My work during the past year has aimed particularly, then, at a definition and analysis of the factors involved in ability in English composition and their measurement.

You will doubtless feel that the natural method, in investigating ability in English composition, would be to use some one of the standard scales for measuring merit in English composition, of which the Hillegas and Willing Scales are types. However, skill in use of such scales is only attained by some considerable practice. And, even then, their use is very time-consuming. Since we had adopted as our motto that we would "Reduce the high cost of testing" — emphasis being particularly on time costs — use of such scales hardly fitted in well with our program. But there were other objections besides this. There is certainly much question as to the accuracy of the measurements yielded by such methods. Our first effort was, then, to investigate the accuracy of the measures yielded by such instruments as the Hillegas and Willing Composition Scales, in measuring the quality of English composition. It is well known that different judges may differ decidedly in the values assigned to compositions, when rating by means of comparison with these scales. I felt, however, that there were elements of unreliability in the method that had not been heretofore adequately taken account of. Especially I wished to know whether children who wrote well on one topic wrote comparatively as well on another topic, whether children who wrote narrative compositions well also wrote good descriptive compositions. I therefore had several students rate various sets of compositions from the same children. The ratings were then correlated. The ratings on narrative and descriptive compositions, or on successive narrative compositions written by a given child, were surprisingly inconsistent. Thus, ratings on one narrative composition correlated with ratings on another narrative composition written two or three days later, in a fifth grade class, with a coefficient of only .15. You feel, perhaps, that the ratings must have been unsatisfactory. But both ratings were made by the same student, using the Willing Scale, and the compositions were procured according to Willing's directions. And the student was conscientious and an experienced user of the scale. A number of similar studies were made. The most extensive piece of work was carried thru with a total of 102 cases, a homogeneous group consisting of all the 8A children at Bedford, Ind.; — Superintendent Montgomery very kindly coöperated in the study. The first set of themes was narrative, the second descriptive. These two sets of compositions were rated very carefully by an experienced English teacher, who had taught in both junior high school and high school. The correlation of her ratings on the narrative and descriptive compositions was only .32.

Such a low correlation surely suggests that use of such scales cannot but yield a result largely inaccurate. We wished to check the results, however, in another way. Three of the teachers, who were well acquainted with the written work of the children, therefore estimated the ability of these children, in written English. These ratings were then averaged; this average was considered the best possible evidence as to the sustained general ability of these children in written expression. The correlation between the ratings on the narrative composition and these estimates was found to be only .29. The correlation between the descriptive composition and the average of the estimates was only a little higher — .33. It surely seems not unreasonable to infer that ratings gained by means of any one of the standard scales for measuring merit in English composition must be considered distinctly unsatisfactory because of these various elements of unreliability.

The question then was as to whether a more efficient means of measuring ability in written composition might not be found. The scales I am presenting to you today are the final result. It was our conclusion, from the results that I have just presented to you, that "ability in English composition" was, very decidedly, *not* a single ability but a complex of abilities. So analysis was desirable. And it was to be expected that a total measure might best be obtained by a combined statement of the results of the analysis.

We felt that the formal elements were the first elements to be investigated. We wished, then, measures of ability in spelling, in punctuation, and in grammar. We also felt that a much neglected element, in both ability in silent reading and ability in written expression, was vocabulary. We therefore wished a test of vocabulary. The final result of our efforts appears in the blanks which are now being passed out to you. The directions and the first five lines of the vocabulary test run as follows:

Name............................ Class or Grade...... Section.......
DIRECTIONS: Below are some statements such as
An oak is a: **tree fish vegetable animal.**

Here the word which correctly defines "oak" is "tree", so "tree" has been underlined. The statement now reads, "An oak is a tree", which is true. In each of the statements below you have four choices for the correct word in just this way. One, and only one, is right. This correct word you are to underline, just as "tree" is underlined in the example. If you cannot be sure, guess. Work as rapidly as you can.

1. A cog is a part of a: **building bridge boat machine.**
2. A hypocrite is one who: **admires favors deceives dominates.**
3. Anything genuine is: **impossible intellectual original copied.**
4. A person who is vaccinated is: **dangerous poor tired immune.**
5. The duty of a jury is to: **plead decide justify abjure.**

The grammar test is organized as follows:

Name.................................Class or Grade... Section.... Age.....
DIRECTIONS: On each of the lines below are four sentences such as:
It isn't true. We was home. How are you? All are here.

One of these four sentences—"We was home"—is grammatically incorrect; this wrong sentence has been underlined.
In each of the lines below there are four such sentences. One, **and only one** is grammatically incorrect. You are to find, in each line, this wrong sentence, and draw a line under it. Work as rapidly as you can. And remember, underline the incorrect sentence. If you are not certain which sentence is incorrect, guess.

1. I wrote to him and her. He spoke to her and me. She stands better than I. Between you and I, it's not so.
2. To whom should it go, you or me? Who did they say was hurt? Who will the book be read by? I know whom he struck.
3. She is far ahead of me. He likes me but not them. I am older than him. You work as well as she.
4. That seems to be he. You and they must decide. We girls are all going. Does that apply to we older boys?
5. Every girl took her seat. Each of us have one. Neither of them is right. Has any one lost his knife?

The nature of the punctuation test can, again, be inferred from the following sample material:

Name.............................. Grade or Class . . . Section . . . . .

DIRECTIONS—READ CAREFULLY: The sentences below lack all punctuation and capitalization except the capital at the beginning of each sentence and the period at the end. You are to supply all further punctuation marks and capitals, changing periods to interrogation marks or exclamation points where necessary. In many of the sentences **more than one** mark or capital is needed; be careful to put in **all** that should be there. Where you cannot be sure of the punctuation, guess. Make all your marks clear and plain; any doubtful marks will be counted against you.

---

1. We have apples peaches plums and grapes.
2. He was satisfied she believed with the report.
3. He went home every saturday in july.
4. The mothers gown and both daughters dresses were white.
5. He hasnt replied because youve been away.

The very brief and compact "Manual of Directions" for giving, scoring, and recording the results of the test appears on the two sides of a single small sheet, $8\frac{1}{2}$ by $5\frac{1}{2}$ inches, as shown below. (These direction sheets are printed on light blue paper so that, even in the pack, they cannot easily be confused with the blanks to be passed out to the children.)

Indiana University Department of Psychology

DIAGNOSTIC TESTS IN ENGLISH COMPOSITION

(Vocabulary, Grammar, and Punctuation)

For Use From Junior High School Thru College

Directions for Giving

Any one of the three tests may be used by itself; for an adequate diagnosis all three tests are needed, however. When all three are to be given it is suggested that the vocabulary test be given first, then the grammar test, and finally the punctuation.

The procedure is the same for each test. First, before distributing the blanks, make certain that all are supplied with pencils (or pens, if any do not have pencils). Then give each pupil a blank **face down**; see to it that no one turns over his blank before he is told to do so. As soon as all are supplied with blanks, read the following directions, **verbatim:**

Now, turn over your blanks. Write on the line at the top of the page your name, your class (or grade), **and your section** (if class is so divided). **Then read the directions at the top of the page, and do what they tell you to do. Work rapidly.**

Do not permit any questions about the test. Immediately after finishing the above directions note the time, and collect the papers at **once** at the end of 5 minutes (Vocabulary), 15 minutes (Grammar) or 10 minutes (Punctuation).

Directions for Scoring

Each test has 40 lines. Each "line" correct counts one point; the highest possible score on each test is thus 40 points. It is suggested that in scoring a check be made against the number of each line correct; total score on each blank is then total number of check marks. Scoring can be most easily done by comparison with a blank on which the correct answers have been marked, conspicuously, in red. Special directions for each test run as follows:

(a) Vocabulary. Each line in which the correct word, **and no other**, has been indicated, counts one point; **any** mark (line under, cross, check) unmistakably indicating the correct word is to be considered satisfactory. In each line there are four choices for the word completing the meaning of the sentence. The correct word is first, second, third, or fourth as indicated by the following numbers (the first number is the number of the list):

1-4, 2-3, 3-3, 4-4, 5-2; 6-4, 7-3, 8-1, 9-2, 10-2;
11-4, 12-1, 13-1, 14-2, 15-1; 16-3, 17-3, 18-2, 19-4, 20-2;
21-1, 22-3, 23-2, 24-4, 25-1; 26-3, 27-1, 28-2, 29-4, 30-4;
31-1, 32-3, 33-2, 34-1, 35-3; 36-4, 37-4, 38-2, 39-1, 40-3;

(b) Grammar. Each list contains four sentences; the sentence to be marked is first, second, third, or fourth according to this same key—and the same rules hold.

(c) Punctuation. The scoring for the punctuation test is given on a separate mimeographed sheet.

RECORD SHEET DIAGNOSTIC TESTS IN ENGLISH COMPOSITION

Grade...... School.................. City..........................

Date............ State........... Examiner.....................

Record: For each test, arrange the blanks in order of total score. (Keep the results from different grades, classes, or sections separate, and record on separate record sheets.) Then tabulate the results in the spaces on the right—if 2 pupils make a score of 29 on the grammer test write "2" in the "Grammar-total" column on the "29" line, etc. Then add together all the figures in the total column and write the sum in the space marked "No cases". Finally, count up from the bottom to the middle (or median) case, and write the number of the line in which this case falls in the space marked median".

To analyze, go through the papers again, making a mark or dot, in the column headed "analysis", for each item correct. Thus if the first paper shows the 39th line of the grammar test correct make a small mark in the "Grammar-analysis" column and in the "39" space. Then count up the total number of marks in each space. The spaces in which there are fewest marks indicate the grammatical errors on which the class is weakest. Drill in "correct English" should emphasize these errors.

Analysis of the punctuation test results should be made in the same way. Analysis of the vocabulary test is not worth while.

The lines of the punctuation and grammar tests have been arranged in ten groups according to the type of problem. The grouping used in the punctuation test is given on the punctuation score sheet. The grouping for the grammar test runs as follows: Lines 1-4 case of pronouns; 5-8 agreement of pronouns; 9.12 agreement of verbs, 13-16 tense and mode; 17-20 confused verbs; 21-24 irregular verbs; 25-28 word order; 29-32 redundancy; 33-36 adjectives and adverbs; 37-40 prepositions and conjunctions. If desired, total passes in each group may next be found. Results on individual lines are, however, most important.

The tabulation sheet is intended only for summary of the results. It is urged that the teacher keep the original blanks and use these as a point of departure in remedial instruction. At least, a record of total scores should be kept in the class book.

Voc. Tot.		Grammar Tot.	Grammar Analysis	Punctuation Tot.	Punctuation Analysis
	40				
	39				
	38				
	37				
	36				
	35				
	34				
	33				
	32				
	31				
	30				
	29				
	28				
	27				
	26				
	25				
	24				
	23				
	22				
	21				
	20				
	19				
	18				
	17				
	16				
	15				
	14				
	13				
	12				
	11				
	10				
	9				
	8				
	7				
	6				
	5				
	4				
	3				
	2				
	1				
	0				
	No.		No.		No.
	Med		Median		Median

Your attention is particularly called to the diagnostic nature of the grammar and punctuation tests. The diagnostic grouping of the lines in the grammar test is given on the record sheet. The diagnostic arrangement of the punctuation test is a little more complex. Sentences 1, 11, 21, and 31 have to do with the use of commas in setting off words. Sentences 2, 12, 22, and 32 involve the use of commas in setting off phrases or clauses. Sentences 3, 13, 23, and 33 require capitals; and so on — the fourth group has to do with possessives; the fifth group, apostrophes and abbreviations; the sixth group, semicolons; the seventh group, question marks and exclamation points; the eighth group, colons; the ninth group, quotations; and tenth group, special use of quotation marks. The vocabulary test is made up simply of words selected at random from a small dictionary. I might say, however, that we have under way a special analytical vocabulary test which aims to indicate comparative knowledge along different lines.

Well — so much for the tests themselves. I think you will grant me that they have an obvious relevancy to the factors involved in ability in written English with which they are supposed to deal. They would appear also of some little analytic value in investigating knowledge of the elements of good usage and, perhaps, in assisting instruction along these lines. As a matter of fact, we have investigated their usefulness in some detail. I shall not worry you with the details of this validation, however. I wish only to give you, in closing the discussion of the tests, one striking bit of result for comparison with the data I have already mentioned dealing with the value of the standard scales for measuring the merit of compositions.

I said a minute ago, that we came to the conclusion that "ability in English composition" was really a complex, involving a number of factors. It was further suggested that a total measure of this complex might best be obtained by analyzing the situation and then combining the results of the analysis. We have attempted to do just this thing.

We took these analytical tests (in a preliminary form) and gave them to the same 8A group from which compositions had already been obtained. We then combined results on the tests, and correlated the total with the pooled teachers' estimate of ability in English composition which I have already mentioned. You may remember that careful ratings of sample narrative and descriptive compositions written by these children correlated .29 and .33 respectively, with these estimates. And now the surprising contrast — our diagnostic tests (in this preliminary form) correlated with the pooled estimate with a coefficient of .72!

So we have some evidence that these tests not only yield a diagnosis of difficulties in English composition; they also combine to give a highly valuable indication of total ability in written English. But I must not dwell upon this point longer. Instead, I wish to indicate to you, in closing, certain practical advantages of these materials. And by way of contrast I wish to mention certain examinations which have recently appeared which, to my mind, are decidedly *impractical.*

A few days ago I was looking over a new examination for measuring general intelligence, for use in the first three grades. The blanks for this scale cost $6 a hundred and stencils for scoring and manual of directions and transportation are extra at that. But the money cost is not the worst of it —

the worst of it is the extraordinary demand on the teachers' time which use of this examination entails. The directions to the children cover 14 pages in the "Manual of Directions"; the general directions to the examiner and the directions for scoring cover 12 pages more. In order to give this examination, these 14 pages must be practically memorized; the manual advises that the directions should not be read, but should be given from notes made on a copy of the blank used by the children, after "careful study" of the printed directions to the children, in the manual. That the examiner must indeed have a mastery of the verbal directions may be realized when the number of other things this unhappy person must do is appreciated. Not only must he say the directions to the pupils. He must illustrate certain tests on the board, occasionally using red, green, and blue chalk. He must point out certain places in the test blank to the children, watching to "make sure that the pupils on his extreme right and left, as well as these in the rear rows, see where he is pointing". He must show the pupils "how to lift the forearm only, leaving the elbow on the desk to avoid fatigue", in raising the pencil between tests, and should "see that all children have their pencils up". He must do this, in fact, 40 times (something of a job, with a restless first grade class!); he must also watch for a hand raised to indicate a broken pencil, and supply new pencils at once. He must observe closely the progress of the class, on most tests, so as to stop work "when from half to three-quarters of the class have finished", tho on one test he is told to call time when half have finished, while in another test the children are to "stop work when three-fourths of the class have finished". In another test he is to proceed to the harder items or not according as "the class as a whole is responding well". Yet we are told that "The series can all be given by the classroom teacher"!

So much as an example of tests difficult of administration. Now a word with regard to tests involving difficulty of scoring. There are examples a-plenty which one might choose. Perhaps the most troublesome type of difficulty, at present, is the test involving elaborate statistical handling, with a recalculation and transmutation of scores, or with methods lacking thoro systematization. Thus, in a recent examination there are 15 tests. In 4 of these, the children are to answer by underlining a word; in 2, they cross out a word; in 10 they write a number. This is not as systematic as it might be. But the real difficulties have not yet been reached. After one has counted the number or words correctly crossed out or underlined, or the number of figures correct, has the final score been obtained? By no means! In 1 test the score is "the number of exercises right minus the number wrong and in case the difference is negative, call it 0". In another it is "the number of spaces correctly filled in, divided by 4"; if a pupil has all 150 spaces correctly filled in, he receives the highest possible score which is "150 divided by 4, or 37½" (apparently ½ and ¼ scores!). In 3 other tests, the correct score is, for the purpose of combining in the final rating, to be divided by 3 — and in 2 other tests, it is to be divided by 2. In another test a "rate" score is obtained by taking "the number that the pupil has marked as indicating the line he was reading when time was called and dividing this number by 4". The number of exercises done correctly makes up a "comprehension" score. However, on this test, the end is not yet. The scorer is told to "note that in the case of examination II, 1 is to be added to the 'comprehension' score and 15 to the 'rate' score, before the scores are transmuted into quotients". And is this all? This is only the beginning. All the scores must now be copied from the 15 inside pages — 8 of which are upside down — on to the front of the booklet, and then added, to give a "total point score".

However, this is not the final form of statement advised. The scorer now adds "43 to the point score and divides by 10; the quotient is the mental age". But a percentage statement is desirable. So the result of this last operation is divided by the child's chronological age (with the help of a table which is kindly provided for this purpose). That is, this is done if the child is not over 14; if he is, the arrangement is that one "add 43 to the pupil's point score and divide this by the average score for his chronological age, plus 43". But for ages above 18, the point score is the same as for 18.

(It is only fair to say that tables are provided for facilitating some of these calculations. The tables cut down the labor somewhat, but hardly reduce the essential complexity of the situation.)

All the calculations described in the last paragraph are for but one-half of the examination. The scorer now goes thru much the same process for the other half and, by the relation of the two halves, obtains some idea of a child's school work in proportion to his ability. But the person who is scoring the examination — presumably the teacher — is not thru with it all yet; the results must be recorded on the class record sheet — which is no small task. And yet the deviser of this scale speaks of his method of handling data as unusually simple!

With such time- and patience-consuming examinations I should like to have you contrast our little "battery" for attacking the problem of "ability in English composition". I would have you note, in the first place, that the blanks are very cheap. The blanks, with directions and all other materials, cost only 30 cents per hundred; blanks for a class of 35 would cost you, then, only about 10 cents. Money costs are thus low. So also are time costs. The tests are very easy, indeed, to give and could be trusted to any of your teachers. All that is necessary is to pass out the blanks, face down, and then, once all are supplied, to tell the pupils to "Turn over the blanks! Read the directions, and do what they tell you to do". Nothing more is required of the examiner, except that he collect the blanks at the end of the alloted time, which is an even number of minutes and thus can be readily kept. Scoring is (with the possible exception of the punctuation test) easy and objective, and interpretation of results not difficult. Total score on all the tests combined is obtained by simply summing the scores on the separate tests.

The tests are, thus, both useful and usable; they are practicable so far as both money and time costs go. I wish to emphasize this point. We are making a special effort this year to "cut the high cost of testing". We feel very decidedly that many of the scales at present on the market are so expensive as to prevent their use by many school systems — so expensive as to limit seriously the further testing program of such schools as may be able to invest in them. We also feel that many of the present tests and scales make demands upon the teacher using them, in the way of time and effort in giving, and particularly in scoring, which are almost intolerable. We feel that tests should not be a burden; they should, rather, make easier the school work; and we feel that the tests should be cheap enough so that use of a variety of test materials may be possible. Thruout the past 4 years, my wife and myself have been the only workers with tests of intelligence who have consistently kept clear of elaborate booklet examinations; no examination which we have issued consists of more than a 4-page folder. The tests which we are presenting you, at this conference, are a step further as regards cheapness and ease in use. The cost of the blanks has been reduced to the point where they should be readily available for anyone. The tests are self-giving; almost no preparation is needed in order to administer them. And the scoring is comparatively easy. Any of you who have had experience with some of the elaborate examinations recently issued will appreciate, I am sure, these features.

Reading Scales for the Second, Third, and Fourth Grades

LUELLA W. PRESSEY

Two years ago at this conference, I presented to you the *first* group scale for measuring general intelligence, devised for use with children in the first three grades. Since that time some two hundred thousand of these scales have been used. Last year I presented our "Scale of Attainment No. 1" for measuring achievement in reading, spelling, and arithmetic in the second grade. This morning I wish to describe, very briefly, an attempt to develop group tests for measuring, and analyzing progress in reading, in the first four grades. Our reading vocabulary test for the first grade has now been in use for almost a year, however. And data regarding this scale has already appeared — in the *Elementary School Journal* for December of last year. So I shall deal primarily with our new tests for the second, third, and fourth grades. And I wish to discuss two points with regard to these tests. I wish, first, to make clear to you our general basis for formulating the tests and analyzing the reading problem in the way we have done, and our evidence that the tests *do* measure reading ability in these grades. I shall then point out, shortly (in the second place) certain practical advantages which these tests have which make them unusually practical and usable.

There are already in existence a few group reading tests that may be used in the second, third, or fourth grade. Prominent among these are the Courtis Silent Reading Test, the Monroe Reading Test, the Haggerty Reading Test, and the new test from the Russell Sage Foundation. All these tests are supposedly useful in the lower grades, but only the Courtis and the Haggerty go as low as the second grade. The form of all these tests, with the exception of the Courtis, is distinctly artificial, however. Thus, one finds such reading material as "Find the tail on the rabbit and make it longer." Such reading material is of an unfamiliar and often somewhat absurd nature. A further objection to the most of these tests is a purely practical one. Thus the Haggerty test is printed in booklet form and is difficult for the children to take. I have given this test and I have found that many of the children would reach the bottom of a page and never start the next one at all, while others would turn over two pages at once and find themselves confronted with such hard reading material that they could make nothing of it. The Ayres test appears upon a single sheet 19 by 12 inches. The sheet is too large to go conveniently on a desk, and if (something you are not supposed to do) you fold the thing down the middle the crease comes across an item and the page is still awkward to handle. Young children often become absolutely bewildered in trying to get control of a sheet of this size. The items of this test are arranged in columns, but there is little to prevent a child from attacking the problems by proceeding across the top row instead of down the first column. In fact, it is perfectly possible to work vertically, horizontally, diagonally, or to execute a chess move of one to the left and two forward, — all of which adds an element of confusion. After consideration of these reading tests, my husband and I decided that we would attempt to devise tests in silent reading which should be truly applicable to children of the second, third, and fourth grades, and tests which would be convenient as well.

I may say, at once, that tho we have called these tests tests of silent reading ability, we have not considered "silent reading ability" an entity, nor have we taken up the topic from the more usual viewpoint. It would be more accurate if I should speak of these tests as tests which endeavored to investigate certain of the most important factors conditioning readiness in assimilative reading. Presumably many factors operate to condition the readiness with which the average adult reads a piece of printed matter. He will read that matter with understanding and rapidity in proportion as he is familiar with the general topic treated, is interested in that topic, knows the meaning of the words which are used in the passage, and applies himself with attention to the reading. "Silent reading ability" may perhaps best be thought of as a composite of such elements, when one is discussing the problem of silent reading ability as that problem appears in the upper grades. Thus in testing the ability of older children to read a certain type of reading matter, tests in interest, vocabulary, application, etc., would be needed. In the first grade, reading tends to be an end in itself; by the time the children reach the seventh or the eighth grade, reading has become a "tool" with which to get lessons in other subjects. The change from reading as an end in itself to reading as a "tool" surely begins in the second, third, and fourth grades. With these younger children, the problem is to be analyzed differently. The two elements which we have thought to be particularly worth investigating were: (*a*) growth in reading vocabulary; and (*b*) the development of speed in silent reading which means primarily (as we saw it) the elimination of oral reading habits — the child reads rapidly, largely in proportion as he has learned to give up the whispering of the words to himself. The two tests which I have to present to you, then, are a test in reading vocabulary and a test in the speed of reading.

The exact nature of the tests can best be understood by an examination of sample sentences and by quotations from the directions. The tests which we give first is called "A Speed Test in Silent Reading". Sample lines appear below:

1. This man is very out poor.
5. She pretty lived with her mother.
10. The women cowards are very brave.
15. He had a strange dream blue.
20. The man comes over the hill house.
25. The war began two time years ago.

The test endeavors to get at the persistence of oral reading habits by investigating the speed with which children can progress down a page of exceedingly simple sentences. The sentences were all selected from second grade readers and contain easy and familiar words. The items of the test are all of approximately equal difficulty. They were selected from a large number of items, as being easily read by most children of these grades. The question, then, was as to how rapidly these sentences could be read by the children, *with understanding of their meaning*. As a check to insure that the sentence was to some extent, at least, understood, the very simple device has been used of putting in each sentence an extra word, which the children are to find and indicate by making a pencil mark around it. The advantages of this device are, I believe, apparent. The children can hardly find the extra word, without sensing to some degree the meaning of the sentence. But they indicate that they understand the meaning by an exceedingly simple operation — by simply marking, in some way, this extra word. No writing is required — no activity which might interfere to any extent with the progress of the reading itself. The whole method is exceedingly simple and grasped by the children very readily indeed; in a total of 181 second grade children, not a single child failed to make some score; only one child scored below 4. Such a fact is striking evidence that the test is readily understood by even the young children — that the activity called for in the tests is by no means so difficult as, in itself, to present a special problem.

The directions, as you will see, are printed on this separate sheet, which is tinted blue so that the directions can be at once distinguished from the children's blanks. The question and answer method of giving the directions is used. The reason for this method is very simple. Anything formal or artificial simply doesn't go with young children. They will not pay attention unless they, themselves, may take a part in what is going on. That you may understand the nature of the directions, I will read the directions for the speed test to you. The blank is first given out with the test itself face down; the back of the test blank, as you will see, contains spaces for the name, age and grade, and examples. The children are told (after the blanks have been distributed): "Write your name on the line marked 'Name'. Then write your age in years right below in the space marked 'Age'. And after that put your grade." They are then told to:

"Look at the first sentence, just below where you wrote your name. The words are, 'I ball see a boy.' Do you all see the sentence? Now, there is an extra word in that sentence, — a word that doesn't belong there. What is the extra word? Who can find it? Yes, 'ball'. Everyone draw a line around 'ball', because it doesn't belong in the sentence. The sentence should be 'I see a boy'; so draw a line around 'ball'.

"Look at the next sentence. What is the extra word? Yes, 'cat'. Draw a line around 'cat' because it doesn't belong there. The sentence should be, 'I can see my book.'

"Look at the next sentence. What is the extra word? Yes, 'ice'. Draw a line around 'ice' to show it doesn't belong there.

"Look at the last sentence. What is the extra word? Yes, 'tail'. Draw a line around 'tail' because it doesn't belong in the sentence.

"And now, everyone attention! On the other side are some more sentences. *Do not turn your papers over yet,* wait till I tell you to! (See that no one turns paper over.) In each sentence, on the other side, there is just *one* extra word. You are to find the extra word and draw a line around it. Be sure to work *just as fast as you can* and get the sentences right. Now, turn over your papers and start!"

The children are given just 4 minutes to work. They are then told: "Turn over your papers again, at once! The side with your name on it should be up." The papers are then collected. The whole procedure takes less than 8 minutes.

The vocabulary test is quite as simple. As you will see, it consists simply of a small sheet on which are printed such questions as:

6. What do you do with a knife? write sew read cut.
9. What is a coach? a stove a chimney a carriage a door.
14. How does a coward feel? afraid happy sad well.
17. What is a pasture? a forest a river a field a lake.
27. What is a companion? a friend an enemy a beast a ship.
36. What is a tempest? a sea a storm a fight a hero.

The blanks are again passed out, face down, the children being warned not to turn over their blanks until they are told to do so. As soon as all the blanks are passed out, the command is: "Write your name anywhere on the blank side. Under your name put your grade." After the children have all finished writing this information, they are told to turn over the blank and it is explained to them (using the same question and answer method and employing the first 4 lines as examples) that they are to find the answer to each question and draw a line around this answer. Any mark indicating the correct word, and no other, is accepted as satisfactory. It should be noticed again that the response of the children is very simple; a single mark suffices for an answer, so difficulty in writing does not operate to affect the score. That the children, as a matter of fact, *do* understand the directions readily is evidenced by the fact that, out of 181 second grade children recently tested, only 1 child made a score below 9.

The test is not a time-limit test. The intention is to give all the time that is necessary. It has been found, however, that most children either finish, or have progressed as far as they can go, by the end of 8 minutes. The papers are, therefore, collected at the end of this time. The words used in the test were all taken from school readers for the second, third, and fourth grades, and are words which occur frequently in the variety of readers which were examined. The lines are arranged in order of difficulty.

It must not be supposed that these tests were the only tests experimented with, in this attempt at an analysis of the reading problem in the first four grades. Two other vocabulary tests were also tried, and one other speed test. A series of paragraphs with questions following, quite similar to the Courtis Silent Reading Tests, were also extensively worked with. One form of this test has been included in the Third Grade Attainment Scale. All these tests were tried, and were then rated for their efficiency in selecting the good and poor readers, by having the teachers, in the rooms in which the experimenting was done, turn in the names of the 5 best and the 5 poorest readers in the room. The scores made by the poorest readers were then marked in in blue on the tabulation sheets (each room being tabulated separately) and the scores of the five best readers in the room in red. The two tests were then selected which marked off most distinctly the two extremes as thus indicated. Consideration was also given to the extent to which the tests supplemented each other, contributing distinctively to a total measure, and aiding in an analysis.

The tests, as thus made up, were next tried in a survey of an entire school system. Superintendent Montgomery of Bedford, Ind., very kindly allowed me to survey his second, third, and fourth grades with these two tests. At the time of testing, each teacher of a third grade class was given a blank upon which she was to make ratings of the "reading ability" of her pupils. This reading ability was further defined as the "ability to comprehend, quickly and easily, the material encountered in the regular school work". These estimates were made on a form modeled after the army officers' rating scale. The estimates were collected on the same day that the tests were given. The teachers were also asked to check the names of those children who showed the persistence of oral reading habits by such evidences as whispering, moving the lips, following the text with the finger, and so forth. After the tests had been scored, they were correlated with the teachers' estimates. There were, in all, 193 pupils in the third grade, all of whom were included in the correlation. The correlation of rating with the vocabulary test was .59; the correlation of rating with the speed test was .61; and the intercorrelation between the tests was .57. A further correlation was made between the ratings of the teachers and the total on the two tests. This coefficient was .65.

Such data at once call for further analysis. Since the combination of the two tests raises the correlation with the ratings, it appears that the two tests supplement each other; as was desired, the two tests seem to measure somewhat different factors in contributing to a total indication of ability in assimilative reading. To analyze such a correlation situation, the method of partial correlations is required. The partial correlation between estimated ability and vocabulary, "speed being constant", was found to be .38. This means that the distinctive contribution of the vocabulary test — the contribution *not* made by the speed test — has a relationship to estimated ability of about this amount. The correlation of the ratings with speed, "vocabulary being constant", was found to be .41. Each test may then be considered to yield information not given by the other, the combination of the two tests informing us, with no slight degree of accuracy, regarding the reading abilities of the children tested.

As a final check on the validity of the tests, the scores of those children indicated by the teachers as having oral reading habits were marked in on the grade distributions of the Speed Test. All these children scored in the lowest 10 percentile group.

In interpreting the results from these two tests, the teacher is guided by the following suggestions, on the directions sheet, which are based on these results:

Interpretation of Results. A high score in Vocabulary and a low score in Speed would indicate that the child (or class) is still the victim of oral reading habits which retard speed. A low score in Vocabulary and a high score in Speed would indicate the need for more advanced reading material. Low scores on both tests would suggest both poor reading habits and small vocabulary. High scores on both tests would suggest the possibility of special advancement. In any case the teacher should use the test scores as a point of departure for intensive study of her pupils.

The last point should be especially emphasized — that the test results are only *starting-points* for further study. Any child making an a-typical score on these tests should be made the object of intensive study along such lines as those suggested above.

So much, then, for the first part of my paper. I have attempted to make clear to you our point of view in formulating the tests, our general reasons for analyzing the reading problem in the way we have done, and our evidence that the tests *do* measure reading ability in the grades dealt with. Now, I wish to say a few words with regard to the second point mentioned at the beginning of the paper, — with regard to certain practical advantages which these tests have, making them unusually practical and usable.

As has already been indicated to you, in my husband's paper, and enforced by our various charts in varied colors, the slogan we have adopted for this year's work has been "Reduce the high cost of testing!" and we have had in mind not only money costs, but, particularly, time costs. I wish to emphasize to you the way in which these tests are an evidence that we have worked true to this motto of ours.

First, regarding money costs. We are selling these blanks at 30 cents a hundred, which means that blanks for a class of 35 will cost you about 10 cents. I should like to contrast this with an examination, "Sigma 1" for grades one to three, issued by the World Book Company, arranged and standardized by Haggerty and Miss Noonan, and used in the Virginia Survey. These blanks cost $5.60 per hundred. In our price, of 30 cents per hundred, is included our little combined sheet of directions for giving and scoring with table for tabulation of results. If you use this more expensive examination it is necessary for you to pay extra for a score sheet. Moreover, in order to get the directions for giving you must buy a small volume, of 64 pages, including directions for two other scales and costing you 35 cents. Further, transportation on this bulky material is by no means a negligible amount. When you contrast our 30 cents with Haggerty's total of between $6 and $7, I think you will grant me that we are not trying to profiteer.

I said, however, that we are particulary interested in time costs. And in this connection I should like to have you consider certain items in the Haggerty Scale. For instance, item number 2 is a picture of a fish and a bird; and the reading matter consists of this command: "Make an eye on the bird!" Those of you who are familiar with first and second grade drawings will grant me, I believe, that it is not always easy to tell whether a first grade child has "obviously intended" a mark to be an eye or not. In another item a pig is shown, and the child is told to "Put a tail on this pig"! We are kindly told, in the scoring directions, that "any kind of tail put on in the right place scores right." However, the situation is not always so clear. For instance, the child is shown a picture of the limb of a tree with 2 birds sitting on it, 1 bird flying, and 4 birds sitting on the ground; the birds are very small, only about a quarter of an inch across. The children are told to "Put a cross over each bird that is on the ground." And in the scoring directions we are told that "Each cross must be above a bird, not on a bird, to score right. Score right if there is a cross above at least 3 of the birds that are on the ground." But in the very next item we have two children hiding in a box, and another little girl who has found them. And the reading matter below tells the pupils to "Put a cross on each child that is hiding. Put two lines under the girl who has found the children." And the scoring directions are "The cross must be *on* each child in the box to score right. Score right if the crosses are on the hair, face, or clothes of the children. Score wrong if cross is also on the girl outside of the box. Lines must be *under* girl outside of the box only to score right." It may surely be questioned whether ability to read with understanding, and ability to follow such exacting and irrational directions are the same thing. But altogether aside from this is the difficulty of scoring such a test, once it has been given. Just think, if you will, of the nice problem of deciding whether each cross, laboriously made by the stubby pencil of a chubby-fingered first or second grade infant, is satisfactorily placed with respect to a quarter-of-an-inch bird! I have scored this test myself, and I assure you that it is "some job". And I should like to have you contrast the scoring problem of this examination with the simple requirement of the tests I just presented to you, the one rule being that "The correct word, and no other, in each line must be marked." These little tests of ours can easily be scored at the rate of three a minute. I doubt if anyone could get thru a Haggerty examination in less than 4 minutes, — not to speak of getting thru it with any peace of mind.

Our little tests are, then, cheap in money, and cheap in the time and effort required to use them. We trust you will be interested in experimenting with these materials.

Standardized Measurements As a Means of Determining Most Effective Class Size

HUBERT G. CHILDS

IN this period of increasing tax burdens for school support, a maximum of educational efficiency with a minimum cost is greatly to be desired. Class size is one of the most important factors in the production of instructional costs. Class size in our elementary schools has been greatly reduced from the numbers of former years, but what evidence have we that our present standards are justifiable? What has determined our present size of class: the burden the teacher can stand or a high type of achievement on the part of the class?

During the past twenty years a number of investigations have been made to determine the relation of class size to efficiency of class results. All but three of these have concerned themselves with classes in the elementary schools.

Methods of Determining Efficiency

Various methods of determining efficiency of results have been employed, namely:

(1) Promotion rate (the per cent of pupils promoted at the end of the term).
(2) Withdrawals (the per cent of pupils withdrawing during the term).
(3) Teachers' marks (the median or average class scores as reported by teachers at the end of the term).
(4) The achievement of first year college students entering from high schools with a majority of large or small classes (based on average grades received in college).
(5) Teachers' estimates of ideal class size.
(6) Standard tests or examinations given in various grades of school work.
(7) Growth during a given period as determined by standard tests or examinations.

Limitations of These Methods

The method of promotion rate is subject to the limitations of variations in teachers' marks and of administrative demands that non-promotions must be limited.

The method of withdrawals takes no account of the variety of causes for withdrawals, but assumes by implication that withdrawals are due to ineffective teaching. This method also ignores the impossibility of securing accurate statements of the causes of withdrawals from pupils and the inadequacy of school records relative to withdrawals.

The method of teachers' marks involves all the inaccuracies due to varying and subjective standards of individual teachers as to what a given score or grade represents in the way of pupil achievement. These variations are marked for the same teacher in a given subject at different times, for different teachers in the same subject, and for teachers in different subjects.

The method of judging efficiency in high school instruction by the grades earned in college involves all the errors of teachers' marks, takes into account only a minor part of all high school pupils, and those the most capable who will succeed reasonably well regardless of the organization and quality of instruction; overlooks the fact that large schools, where large classes are in highest proportion, generally have the more experienced and highly selected teaching staffs; and also overlooks the fact that pupils in larger schools have learned to look out for themselves in an impersonal and mass atmosphere while pupils from the smaller schools are under closer supervision and may fail to work as successfully the first year in college where the close supervision does not exist.

The method of teachers' estimates of ideal class size is subject to the limitations of general impressions, the influence of the class size teachers are most accustomed to (they generally estimate the ideal class size a little smaller than the size of the classes they are teaching, except in the case of very small classes), and the influence of the inadequacies of their teaching equipment.

The method of using standard tests at a given time as a measure of achievement fails to take into account the probable inadequacy of the tests to include the essential features of the subject as emphasized by the different teachers whose classes are tested, and overlooks the influence of conditions of previous years, particularly differences of pupils in achievement when first entering the classes of the teacher whose classes are later tested. This method also has the limitations that in many of our subjects no standard tests have, as yet, been devised.

The method of measuring growth during a given period by standard tests or examinations before and after the training period is an improvement over the single application of the standard test, but otherwise is subject to all the limitations noted above for the method of using standard tests. With high school classes beginning a new subject, as algebra, a single measure at the end of a given period represents the growth during the period.

Results and Conclusions of the Investigations

1. By the method of promotion rate.
    - *a.* Cornman, 1909. Class size has little or no influence on promotion rate (in the elementary schools).
    - *b.* Boyer, 1913. There is no apparent relation between class size and promotion rate (in the elementary school).
    - *c.* Bachman, 1913. The promotion rate is nearly the same in classes under 35 and over 50 and there is no warrant for changes in the policy of assignment of class sizes (in the elementary school).
    - *d.* Elliott, 1914. There is no relation between class size and promotion rate (in the elementary school).
    - *e.* Harlan, 1915. Class size is not an important factor in promotion rate, altho small classes make a slightly better showing (in elementary schools).

2. By the method of withdrawals.
    a. Harlan, 1915. The larger the class the higher the percentage of permanent withdrawals (in elementary schools).
    b. Neet, 1920. For first year high school classes in English, Latin, French, algebra, history, and sciences the per cent of withdrawals is greatest (10.7 per cent) in the smallest classes (those having 15 pupils or fewer) and is nearly the same for classes of size 16 to 20, 21 to 25, and 26 or more (7.1 per cent, 7.3 per cent, and 7.7 per cent respectively).
3. By the method of teachers' marks.
    a. Neet, 1920 (20 Indiana high schools). The most efficient class size varies for different subjects and different years of work. On the whole, the middle range classes, 16 to 25, show slightly better results. In the first year high school subjects of English, Latin, French, algebra, history, and sciences classes of 26 or more achieved higher average grades than classes of 15 or fewer, and classes of over 20 higher than classes of under 20. The highest average grades in English 1 and History 1 were made by the classes of 15 or fewer, but in French 1 and Sciences 1 the highest averages were made by classes of 26 or more. For Latin 1 and Algebra 1 the results were substantially the same for classes of 16 to 20, 21 to 25, and 26 or more, the lowest average score being in classes of 15 or fewer.

TABLE 1. FIRST YEAR RESULTS IN ENGLISH, LATIN, FRENCH, ALGEBRA, HISTORY, SCIENCES, BASED ON TEACHERS' MARKS

(All subjects combined)

CLASS SIZE	NO. OF CLASSES	AVERAGE SCORE	AVERAGE PER CENT OF FAILURE	AVERAGE PER CENT WITHDRAWN
Under 16....	55	82.0	13.9	10.7
16 to 20..	103	81.5	14.7	7.1
21 to 25..	133	82.9	11.7	7.3
Over 25....	107	83.7	13.5	7.7

TABLE 2. FIRST YEAR RESULTS IN ENGLISH, LATIN, FRENCH, ALGEBRA, HISTORY, SCIENCES, BASED ON TEACHERS' MARKS

Average Scores

CLASS SIZE	ENGLISH 1A	LATIN 1A	FRENCH 1A	ALGEBRA 1A	HISTORY 1	SCIENCES 1
Under 16..	87.3	77.0	79.5	79.3	84.0	83.4
16 to 20...........	84.8	82.0	83.0	85.5	81.0	83.1
21 to 25.	85.0	82.8	83.3	82.0	82.0	84.3
Over 25	83.5	82.0	88.0	84.0	82.0	86.3

Best results—
    Smallest classes: English and history.
    Medium to large classes: Latin and algebra.
    Largest classes: French and sciences.
Poorest results—
    Smallest classes: Latin, French, algebra.
    Medium small classes: English, history, sciences.
    Medium large classes: no subjects.
    Largest classes: no subject.

4. By the achievement of first year college students entering from certain type high schools.
    - *a.* Pittenger, 1915. The best results were achieved in college by students who came from high schools having class enrollments of from 21 to 30, and the poorest results by those coming from high schools having class enrollments of under 20.*
5. By the teachers' estimates of the ideal class size.
    - *a.* Cook, 1920. Classes of 20 to 25 are pronounced as ideal in the majority of high school subjects. Variations are indicated for different subjects and for different phases of the same subject.
    - *b.* Neet, 1920 (based on replies of all teachers in 38 Indiana high schools). "The consensus of opinion of high school teachers seems to be that in every subject the highest efficiency is attained in classes of 15 to 20, inclusive." About 70 per cent of all replies are in this range, but 4 per cent are in range 10 to 15 and 7 per cent each in ranges 20 to 25 and 25.

      "Middle sized classes are favored because of easier discipline, more individual attention, greater opportunity to recite, more inspiration."

      The ideal class size indicated above averages from 3 to 5 smaller than the average sized classes in actual practice in the schools reporting.
6. By the method of using standard tests to measure achievement at a given time.
    - *a.* Rice, 1907. In arithmetic and language there appears to be no relation between class size and results (in the elementary school).
    - *b.* Elliott, 1914. No group possesses superiority because of class size (in the elementary school).
    - *c.* Harlan, 1915. There is no advantage for the small class. The median results (Courtis tests) fell in the median sized group (in the elementary school).
    - *d.* Neet, 1920 (data from 16 schools). The highest average score in first term algebra, as determined by the Rugg-Clark tests, was 2.23 for the group of classes of size 16 to 20 and lowest, 1.87, for classes of 26 and more. The average scores for classes of under 16 and 21 to 25 were 2.05 and 2.07 respectively. Median scores are 1.96, 2.08, 2.065, 1.85 respectively. So far as this method has validity and so far as concerns first term algebra the most effective results are secured in classes of the range 16 to 25, the results being materially better than in classes of 26 or more.

      "The average number of rights per group is highest in group 16 to 20, next highest in group 21 to 25, and lowest in group 26 to 30." The per cent of accuracy is about the same for all groups up to 25, but is materially less for the 26 to 30 group.

Miss Neet states the following conclusions as a summary of the findings in the use of the algebra tests to measure efficiency:

---

*This is more a comparison of results achieved by students coming from large and small schools and involves many important factors other than class size.

(1) "The most efficient sized class in first term algebra falls in group 16 to 20.
(2) The least efficient sized class in first term algebra falls in group 26 to 30.
(3) The most efficient class size varies somewhat for different schools.
(4) The most efficient class size varies for different teachers.
(5) Teaching experience in mathematics and training in mathematics by the teacher are more important factors in determining the efficiency of a class than the size of the class."

TABLE 3. FIRST TERM RESULTS IN ALGEBRA AS MEASURED BY THE RUGG CLARK TESTS (FIRST 8 TESTS)—MISS NEET'S INVESTIGATION
Class Scores

UNDER 16	16 TO 20	21 TO 25	OVER 25
1.59	1.33	97	1.21
1.90	1.56	1.29	1.31
2.02	1.64	1.64	1.33
2.70	1.77	1.66	1.34
	2.08	1.85	1.49
	2.89	1.88	1.59
	2.90	1.96	1.64
	2.90	1.99	1.72
	3.12	2.00	1.74
		2.00	1.84
		2.01	1.86
		2.12	1.86
		2.18	1.98
		2.20	2.17
		2.33	2.20
		2.42	2.29
		2.59	2.43
		2.62	2.65
		2.75	2.67
		2.87	2.89
		2.96	
		3.28	
Average 2.055	2.233	2.07	1.877
Median 1.96	2.08	2.065	1.85

Class scores were determined from the formula

$$\frac{\text{Class rights}}{\text{Median rights}} + \frac{\text{class accuracy}}{\text{median accuracy}}$$

7. By the method of measuring growth during a given period of time.
   - *a.* Elliott, 1914. There is no correlation between class size and attainment (in the elementary school.)
   - *b.* Breed and McCarthy, 1916. The larger the class the better the results up to a class size of 45 (in spelling in grades 3 to 8).
   - *c.* Neet, 1920. (See previous heading. The results may properly be considered here as measuring achievement in a new subject at the end of a half year as equivalent to determining growth during a given period.)

Taking into account certain serious limitations in the methods of investigation, the foregoing statements indicate that for the elementary school no appreciable differences in achievement or promotion rate are noticeable for classes of varying sizes up to 45 or 50. These numbers probably convey a wrong impression, as quite generally in elementary schools classes of more than 30 are divided into two sections for recitation purposes, one section being engaged in study while the other is reciting.

Miss Neet's use of standard algebra tests to determine growth in achievement in classes of different sizes is the most accurate and reliable investigation in the high school bearing upon the influence of class size. The findings apply to the formal processes of algebra only, and similar methods applied to other lines of instruction might yield different results. Also it would be desirable to have measures of the general mental ability of all classes tested as mental ability is an important factor in achievement.

An important limitation to all these methods as applied in previous investigations has been the failure to analyze critically the total situation to determine the factors that condition teaching efficiency, and then to so control the experimental situation that results may be definitely referred to the causal factor.

Briefly, some of the more important factors conditioning teaching efficiency are:

1. The native ability of the members of the class.
2. The teachers, including:
    a. Personality.
    b. Teaching skill as evidenced by:
        (1) Training, particularly in the subject taught.
        (2) Teaching experience, particularly in the subject taught.
3. Nature of the subject or of different phases of the subject.
4. Methods used.
5. Class size.
6. Length of time the class has studied the subject.
7. Method of determining the efficiency score.

Further progress in solving this problem must depend on a more careful analysis of the total situation, on a more careful definition and control of experimental conditions and factors. Both pupils and teachers must be carefully selected, results must be considered separately for different grades and subjects and not for all grades and subjects lumped together promiscuously. The measure of achievement must be such that it measures what it purports to measure and can be applied uniformly for all classes tested. Achievement should include not only memoriter knowledge and skills, but also habits of thought, emotional responses, and mental and moral attitudes for whose acquisition intimate contact with strong personalities has been thought necessary.

No single method of attack will be likely to lead to the complete solution of the problem, but contributory evidence from all methods refined and widely applied may enable us to reach conclusions which shall lead to improvement in our present mode of class grouping for both educational and economic ends.

Diagnostic Test in Modern European History

CHARLES G. VANNEST

INDIANA UNIVERSITY maintains, in connection with the Bloomington high school, a practice school for teachers. It is our plan in the history department to have the practice teachers give the various standardized tests in the different fields of history. In looking over the field in modern European history we found no test, so we proceeded to make one. As far as we know there is no standardized test in modern European history.

The Diagnostic Test in Modern European History consists of five parts: Tests A, B, C, D, and E. Test A is a test for time sense; test B for place sense; test C for evaluation of facts; test D for thought; test E for information. Tests A, B, and C each consists of 3 parts. In part 1 of test A the names of 10 personages noted in modern European history are given, arranged in no fixed order. The pupil is asked to place a check mark after any name he has not had in his history work and then to indicate the time order in which the personages lived by placing a figure one (1) before the name of the person who lived first, a figure two (2) before the name of the person who lived next, and so on thru the entire list. Part 2 of test A contains the names of 10 important events in modern European history. The pupil is directed to do here the same as he did in part one of test A. Part 3 of test A contains the names of divisions or periods in modern European history. The pupil is given the same directions here as in parts 1 and 2 of test A.

Part 1 of test B contains the names of 10 personages noted in modern European history. The pupil is asked to place a check mark after any name he has not had in his history work. He is then to underscore the word which names the nationality of each man. Part 2 of test B contains the names of 10 noted European events. The pupil is given the same direction here as in part 1 of test B. Part 3 of test B contains the names of 10 important modern European battles. The same directions are given here as in parts 1 and 2 of test B.

Part 1 of test C contains the names of 10 important facts taken from modern English history. The pupil is asked to place a cross mark after any fact he has not had in his history work. He is then to place a check mark before the names of the 3 events that have been of the greatest importance in the political development of England. Part 2 of test C contains the names of 10 important facts taken from modern French history. As in part 1 of test C, the pupil is to place a cross mark after any fact he has not had in his history work. He is then asked to place a check mark before the names of the 3 events which have been of the greatest importance in the religious development of France. Part 3 of test C contains the names of 10 important facts taken from modern German history. As in parts 1 and 2 of test C, the pupil is to place a cross mark after any fact he has not had in his history work. He is then to place a check mark before the names of the

3 events that have been of the greatest importance in the industrial and economic development of Germany.

Test D contains 10 thought questions of varying degrees of difficulty. A blank space is left following each question in which the answer is to be written. The following is question number 7 of this list:

The Austro-Hungarian Empire was composed of an extraordinary mixture of peoples. In Austria-Hungary were found Germans, Czechs, Poles, Croatians, Italians, Slovenes, Serbs, Ruthenes, Slovaks, and Roumanians. In general each of these people lived adjoining their fatherland, i.e., the Italians in Austria-Hungary lived in territory adjacent to Italy, the Roumanians in territory adjacent to Roumania, etc. These people were not satisfied at all with the Austro-Hungarian rule and were anxious to be annexed to their native countries.

A. Why was the Austro-Hungarian government deeply concerned with what went on in a political way in Germany, Italy, Roumania, Serbia, and in all the countries about her?

B. When Austria-Hungary was defeated in the World War what would you expect would happen to her different territories?

Seven of these 10 thought questions deal with very recent events, most of them leading up to or growing out of the World War.

Test E contains 6 questions, apparently of an informational character. Following the questions, blank spaces are left for the answers. Question 1 in this list is as follows: "What are the 'natural' boundaries of France?" Question 6 is as follows: "What European nations secured possessions in Africa in the nineteenth century?"

Accompanying the test is the following questionnaire to be filled in by the teacher whose pupils have taken the test:

1. Name of city..
2. Population of city..
3. Name of teacher whose pupils have taken the test.................
   ....
4. Teacher's age..
5. Sex...
6. Number years teaching experience................................
7. Number years teaching experience in history.......................
8. Number years teaching experience in modern European history......
   ....
9. Number years teaching experience in present school system.........
   ....
10. Has teacher taken a course in the methods of teaching history?......
    .... Where? .... When?..............
11. Has teacher taken a course in practice or supervised teaching in history?
    .... Where?.................. When?...........
12. Is teacher a high school graduate?...............................
13. Years of normal or college work done by teacher....................
14. College or university degrees held by teacher.......................
15. Date this test was given..

The test is primarily a diagnostic test. It aims to find out as much about the teacher as about the pupil. The teacher's answers to the questions in the questionnaire may explain the pupils' answers in the test papers. If an entire class, or a considerable portion of an entire class, places a cross mark after the name of Lloyd-George in part 1 of test A, we may feel quite sure that the class has not spent much time upon really recent modern European history. If a class states that it has not had the Industrial Revolution in its history work but returns a high mark on the list of the 10 European battles, we are to infer, to some degree at least, that the teacher of that class needs to change her viewpoint of teaching. If the class cannot answer the questions stated above — questions 1 and 6 in test E — we feel quite sure that the teacher has overlooked two very important factors closely connected with the events of the past few years.

In each of the 3 parts of tests A, B, and C the pupil is asked to place a cross mark after any name, event, or fact that he has not had in his history work. By this means we are enabled to tell just what part of the history work has received no emphasis.

By the answers returned in test D, we are enabled to tell whether or not the teacher has been placing emphasis upon recent history, as 7 of the 10 questions deal with very recent events.

In the returns from each school, the pupils' answers are read and scored in connection with the reading and scoring of the questions in the questionnaire answered by the teacher of that school.

The following is the test. It is based wholly upon Robinson and Beard's *Outlines of European History*, Part II, the textbook that is used in the high schools in Indiana. The test has not been standardized and is in a tentative form only. It is now being given in about 50 high schools, large and small, in all parts of the state of Indiana. The results from these schools will render a partial standardization possible.

As to form, there is little that is new in the test. Old forms have simply been transferred to a new field. In this matter the author is indebted to M. J. Van Wagenen, of the University of Minnesota, and Arvil S. Barr, of Evansville College. Valuable suggestions and criticisms have been received from Professors Hubert G. Childs and William T. Morgan, of Indiana University, and from Dean William S. Gray and Professors A. C. McLaughlin, Rolla M. Tryon, and Ferdinand Schevill, of the University of Chicago.

Diagnostic Test in Modern European History

Charles Garrett Vannest, Indiana University

BUREAU OF COÖPERATIVE RESEARCH, Indiana University
Bloomington, Indiana

Name.............................Race.............................
Sex.............................Age.............................
School.............................City.............................
Year in High School.............................Date.............................
Teacher...

TEST A.—TIME SENSE

I. The following list contains the names of ten personages noted in Modern European History.

1. After reading the list through carefully, place a cross mark (x) after any name you have not had in your history work.
2. Indicate the time order in which the personages lived by placing a figure one (1) before the name of the person who lived first, a figure two (2) before the name of the person who lived second, a figure three (3) before the name of the person who lived third and so on through the entire list.

a. ( ) Frederick the Great.
b. ( ) Metternich.
c. ( ) Bismarck.
d. ( ) Lloyd-George.
e. ( ) Louis XIV.
f. ( ) Cromwell.
g. ( ) Napoleon Bonaparte.
h. ( ) Cavour.
i. ( ) Gladstone.
j. ( ) Elizabeth.

II. The following list contains the names of ten important events in Modern European History.

1. After reading the list through carefully, place a cross mark (x) after any event you have not had in your history work.
2. Indicate the time order in which the events happened by placing a figure one (1) before the event that happened first, a figure two (2) before the event that happened second, a figure three (3) before the event that happened third, and so on through the entire list.

a. ( ) The Congress of Vienna.
b. ( ) The Russo-Japanese War.
c. ( ) The Glorious Revolution.
d. ( ) The French Revolution.
e. ( ) The Unification of Italy.
f. ( ) The Treaty of Utrecht.
g. ( ) The Franco-Prussian War.
h. ( ) The Boer War.
i. ( ) The Industrial Revolution.
j. ( ) The Expedition of the Spanish Armada.

III. The following list contains the names of the main divisions or periods in Modern European History.

1. After reading the list through carefully, place a cross mark (x) after any division you have not had in your history work.
2. Indicate the time order in which the periods come by placing a figure one (1) before the division which comes first, a figure two (2) before the division which comes second, a figure three (3) before the division which comes third, and so on through the entire list.
    - *a.* ( ) The French Revolution and Napoleon Bonaparte.
    - *b.* ( ) Unification of European States.
    - *c.* ( ) Absolutism and the Rise of the Democratic Spirit.
    - *d.* ( ) Mercantilism and the Struggle for Colonial Supremacy.
    - *e.* ( ) The World War.
    - *f.* ( ) Social Democracy and Territorial and Commercial Rivalry.
    - *g.* ( ) Reaction, Revolution, and the Rise of the Democratic Movement.

TEST B.—PLACE SENSE

I. The following list contains the names of ten personages noted in Modern European History.

1. After reading the list through carefully, place a cross mark (x) after any name you have not had in your history work.
2. Underscore the word which names the nationality of each man.
    - *a.* John Hampden was an English, German, French, Italian squire.
    - *b.* Peter the Great was a French, Prussian, Russian, Italian ruler.
    - *c.* Luther was a Russian, English, French, German reformer.
    - *d.* William Pitt was a French, Italian, English, Prussian statesman.
    - *e.* Voltaire was a Prussian, Russian, French, Italian, Spanish writer.
    - *f.* Newton was an Italian, Spanish, Prussian, English, French scientist.
    - *g.* Turgot was an English, Prussian, Russian, Spanish, French economist.
    - *h.* Stein was a French, Spanish, Prussian, Russian, English statesman.
    - *i.* Talleyrand was an English, French, Prussian, Italian, Russian statesman.
    - *j.* Garibaldi was a Russian, Prussian, French, Italian, Spanish soldier.

II. The following list contains the names of ten noted Modern European events.

1. After reading the list through carefully, place a cross mark (x) after any event you have not had in your history work.
2. Underscore the name of the country in which each event took place.
    - *a.* The Petition of Rights was a French, Prussian, English, Italian, Spanish measure.
    - *b.* The Estates General was an English, Italian, Spanish, French, Russian governing body.
    - *c.* The Declaration of the Rights of Man was a Prussian, Italian, Spanish, French, English measure.
    - *d.* The Reign of Terror took place in England, France, Prussia, Russia, Italy, Spain.
    - *e.* The Puritan Revolution took place in Spain, Italy, Portugal, France, England, Prussia.

  - *f.* Spinning and weaving machines were first invented in Spain, France, England, Prussia, Italy.
  - *g.* The Zollverein (customs union) was formed in England, France, Prussia, Spain, Italy.
  - *h.* The Dreyfus Affair occurred in England, France, Germany, Spain, Italy.
  - *i.* The Reform Bill of 1832 was a French, Spanish, English, Italian, Prussian measure.
  - *j.* The three-class system of voting was used in England, France, Prussia, Spain, Italy.

III. The following list contains the names of ten important Modern European battles.
  1. After reading the list through carefully, place a cross mark (x) after the name of any battle you have not had in your history work.
  2. Underscore the name of the country in which the battle took place.
     - *a.* The battle of Marengo was fought in Italy, France, England, Prussia.
     - *b.* The battle of Austerlitz was fought in Prussia, Austria, France, England, Italy.
     - *c.* The battle of Lützen was fought in Austria, Italy, Spain, France, Prussia, England.
     - *d.* The battle of Waterloo was fought in England, France, Prussia, Belgium, Holland.
     - *e.* The battle of Sadowa was fought in Prussia, England, Austria, Italy, Russia.
     - *f.* The battle of Sedan was fought in France, England, Russia, Italy, Austria, Prussia.
     - *g.* The battle of Sebastopol was fought in England, Prussia, Russia, Austria, France.
     - *h.* The battle of the Marne was fought in Italy, Spain, Russia, France, Belgium.
     - *i.* The battle of the Pyramids was fought in Spain, France, Portugal, Italy, Egypt, Asia Minor.
     - *j.* The battle of the Somme was fought in Germany, France, England, Spain, Portugal.

TEST C.—EVALUATION OF FACTS

I. The following list contains the names of ten important facts taken from Modern English History.
  1. After reading the list through carefully, place a cross mark (x) after any fact you have not had in your history work.
  2. Place a check mark (v) before the names of the three events that have been of the greatest importance in the political development of England.
     - *a.* ( ) The theory of the "divine right of kings".
     - *b.* ( ) The Separatists leave England and go to Holland.
     - *c.* ( ) The Conventicle Act.
     - *d.* ( ) The Act of Settlement.
     - *e.* ( ) The Reform Bill of 1867.
     - *f.* ( ) The Settlement of Australia.
     - *g.* ( ) The Boer War.
     - *h.* ( ) King James's translation of the Bible.
     - *i.* ( ) Pride's Purge.
     - *j.* ( ) The Habeas Corpus Act.

II. The following list contains the names of important facts taken from Modern French History.
  1. After reading the list through carefully, place a cross mark (x) after any fact that you have not had in your history work.

2. Place a check mark (v) before the names of the three events that have been of the greatest importance in the religious development of France.
    - *a.* ( ) Louis XIV's invasion of the Netherlands.
    - *b.* ( ) Revocation of the Edict of Nantes.
    - *c.* ( ) Reforms of Calonne.
    - *d.* ( ) The Franco-Prussian War.
    - *e.* ( ) The fall of the Bastille.
    - *f.* ( ) The Concordat of 1801.
    - *g.* ( ) The Code Napoleon.
    - *h.* ( ) Execution of Louis XVI.
    - *i.* ( ) Separation of Church and State in 1905.
    - *j.* ( ) France recovers Alsace-Lorraine.

III. The following list contains the names of important facts taken from Modern German History.

1. After reading the list through carefully, place a cross mark (x) after any fact you have not had in your history work.
2. Place a check mark (v) before the names of the three events that have been of the greatest importance in the industrial and economic development of Germany.
    - *a.* ( ) The Zollverein (the customs union).
    - *b.* ( ) Bismarck Prussianizes Germany.
    - *c.* ( ) Formation of the North German Federation.
    - *d.* ( ) Bismarck makes terms with Catholic party.
    - *e.* ( ) Adoption of a protective tariff system (1879.)
    - *f.* ( ) The growth of machine industry.
    - *g.* ( ) Formation of the Triple Alliance.
    - *h.* ( ) Resignation of Bismarck.
    - *i.* ( ) Germany secures Alsace-Lorraine (1871).
    - *j.* ( ) Germany secures African Colonies (1884-1890).

TEST D.—THOUGHT

I. Below you will find ten questions. Answer each question in the space below it.

1. King James II desired to re-establish Catholicism in England. He had married, as his second wife, a Catholic. A son was born to them. James's daughter, Mary, by his first wife, had married a Protestant, William III, the head of the United Netherlands. James also insisted on making laws without the consent of Parliament. The English nation did not want Catholicism re-established and did not want the king to rule without the consent of Parliament.
    - *a.* When the English people invited William and Mary from Holland to become their rulers, what did Parliament insist upon before making them king and queen of England?

2. By the Act of Settlement, George I became king of England, 1714. George I was a German and was ruler of Hanover in Germany. When he became king of England, George I was unable to speak a word of English:
    - *a.* Why was a cabinet organized to help George I rule England?
    - *b.* What line of kings ascended the English throne with George I of Hanover?

3. In the eighteenth century, the government of Poland was the worst imaginable. The Polish king could not declare war, make peace, impose taxes, or pass any law without the consent of the diet. The diet was composed of nobles and any one of the nobles could prevent any measure from becoming law by his veto. Russia, Prussia, and Austria were Poland's neighbors and wanted

to take her territory from her. Whenever the Poles attempted to strengthen their government, Russia, Prussia, and Austria would interfere and prevent them.

   a. Why did Russia, Prussia, and Austria interfere in Polish affairs and help keep the Polish government weak?
   b. What do you think Russia, Prussia, and Austria would eventually do with the Polish territory?

4. England is about the size of the state of Illinois. The population of England is about 40,000,000. On account of its small area, England is unable to raise enough food to support her population. England's chief occupation is manufacturing. England is unable to use all the goods she manufactures.
   a. How do you account for England's colonial empire?
   b. Why does England insist on having the largest navy in the world?

5. Belgium is a low, level plain. To the south of Belgium the country is hilly and mountainous. In this region the Swiss and French nations have heavy fortifications along their boundaries.
   a. Why did the Germans in the World War cross Belgium in their attempt to reach France?

6. (*a*) In 1767, Hargreaves, an English spinner, invented the spinning jenny. In 1768, Arkwright patented a device for drawing out threads by means of rollers. In 1779, Crompton invented the "mule", a combination of Hargreaves' spinning jenny and Arkwright's roller machine. These spinning and weaving machines made possible an enormous output of thread and yarn. Up to this time there was no way to separate the seeds from the cotton except by hand and a worker could only clean about five or six pounds a day working in this way.
   a. Why was the cotton-gin invented?
   b. Why was the cotton-gin invented in the United States?

   (*b*) By the use of the cotton-gin, a slave could take the seeds out of over a thousand pounds of cotton a day, instead of five or six pounds as by hand method.
   a. What effect did the cotton-gin have upon slavery in the United States?

7. The Austro-Hungarian Empire was composed of an extraordinary mixture of peoples. In Austria-Hungary were found Germans, Czechs, Poles, Croatians, Italians, Slovenes, Serbs, Ruthenes, Slovaks, and Roumanians. In general each of these people lived adjoining their fatherland, i.e., the Italians in Austria-Hungary lived in territory adjacent to Italy, the Roumanians in territory adjacent to Roumania, etc. These peoples were not satisfied at all with the Austro-Hungarian rule and were anxious to be annexed to their native countries.

   a. Why was the Austro-Hungarian government deeply concerned with what went on in a political way in Germany, Italy, Roumania, Serbia, and in all the countries about her?
   b. When Austria-Hungary was defeated in the World War what would you expect would happen to her different territories?

8. Turkey has always been a relatively weak nation. Constantinople, the capital of Turkey, has a very important strategic location, controlling the entrance to and the outlet of the Black Sea. It also controls the bridgehead, Asia Minor, through which overland commerce is carried from Asia to Europe and vice versa. Eng-

land is much interested as Asia Minor is on the road to her possession, India. Germany has been developing her commerce with the East, especially in the Tigris-Euphrates region. Russia, the great Slav nation and "big brother" to the Slav states of the Balkan peninsula, has been a strong growing nation, but has been handicapped on account of an outlet to the sea, having never had an ice-free port. Russia has coveted the possession of Constantinople and would have seized it, after defeating Turkey in war, had not England, Austria, and Germany prevented her.

a. Why did England, Austria, and Germany prevent Russia from taking and controlling Constantinople?

9. Belgium is a small state just across the channel east from England. It is, in fact, a buffer state between England and any strong power that might arise on the continent of Europe. England has always insisted that the territorial integrity of Belgium be observed by the European nations. As long as this is done England may feel pretty safe from attack from the east. Ireland lies about as far west of England as Belgium lies east. Ireland is a part of the British Isles but is now endeavoring to secure its independence.

   a. What, perhaps, is one reason why England will not give Ireland its freedom?

10. The east coast of Italy is very regular and has no good ports or harbors. Across the Adriatic Sea from Italy, the Dalmatian coast line is very rough and irregular and affords excellent places for harbors and fortifications. The Italians have frequently stated that on account of the smoothness and regularity of their east coast their country is exposed to attack from that quarter.

    a. Why did the Italian government demand the Dalmatian coast line as Italy's reward for its efforts in the World War?

TEST E.—INFORMATIONAL

I. Below you will find six questions. Answer each question in the blank space below or following it.

1. What are the "natural" boundaries of France?

2. Mention the names of two Spanish explorers in North America.
Two English explorers.
Two French explorers.

3. Name four great benevolent despots of the eighteenth century.

4. Identify the following personages. In each case use but two words in your answer, giving the nationality of the man and the thing for which he is noted.

   a. Adam Smith
   b. Karl Marx
   c. Sir Walter Scott
   d. Rousseau
   e. John Wesley
   f. Gladstone
   g. Darwin
   h. Robespierre
   i. Duke of Wellington
   j. Foch

5. Identify the following dates by telling what important event happened on each of them:
   a. 1832
   b. 1815
   c. 1914
   d. 1648
   e. 1878
   f. 1763
   g. 1871
   h. 1789
   i. 1823
   j. 1848

6. What European nations secured possessions in Africa in the nineteenth century?

School Boards in Indiana

Willis A. Holman

This study was begun in 1917. It was hoped that a more definite knowledge of the personnel of school boards, the fitness of members for the work, what they are doing as board members, and how much power they are giving their school administrators might be obtained. This has been accomplished.

The questionnaire method for gathering data was used. The questionnaire was approved by the Dean of the School of Education of Indiana University. These interrogatories were sent to 228 superintendents of the city, town, and joint schools — that being the number listed in the official directory of 1916-17 where commissioned schools were maintained — 78 cities, 135 towns, 15 joint corporations.

One hundred seventeen systems replied in such a manner that material could be used as a whole or in part. Reports were received from 70 per cent of cities, 30 per cent of towns, and 100 per cent of joint corporations. One wonders just why the smallest per cent of returns was received from the towns. The census report of 1910 was used in tabulating the population of towns and cities. The maximum size of towns entering into a joint corporation was 3,500. There are 35 other towns and 10 cities falling below this population. If consolidation of schools makes for efficiency in school matters, here are 45 of the larger centers that may well unite with the rural communities of which they are a part and make a larger school unit.

Sex

Concerning the personnel of school boards in cities, it was found that 90.8 per cent of members were men and 5 per cent were women. In 4.2 per cent of cases sex was not reported. In towns 100 per cent were men. In joint corporations 96.6 per cent were men and 3.4 per cent were women.

Age

The investigation concerning the age of board members revealed the following facts: The minimum age for members in cities was 27, in towns 28, and joint schools 25; there was only 1 age at 25, and there were only 8 in the 3 groups under 30 years. The recent session of the General Assembly fixed the minimum age limit for certain city school corporations at 25. The minimum age for these cities is already 27. The maximum age given in same order was 79, 70, and 65. The median ages were 45.58, 45, and 44.3 respectively. This makes an average median age of 45 years. Certainly no one can complain that members of school boards of Indiana are unfit because of their youth.

Residence

The next question with which we dealt was that of residence in the community. There was an 83 per cent report on this item for the cities reporting, 90 per cent for towns, 78 per cent for the joint corporations. The minimum years for the 3 types of boards mentioned were 5, 2, 2 respectively. The last session of the General Assembly passed a law establishing a residence qualification for school board members of cities of a certain size, fixing the minimum at 3 years. This report shows that the minimum residence for those cities was 5 years. The maximum years of residence was found to be 79, 70, and 60; the median years of residence, 40.04, 30, and 40 years respectively, and the average median number of years was found to be 36.68. In 5 joint and town corporations appointments were made of persons who had lived in the community under 5 years. This qualification also seems to be taking care of itself very well.

Occupation

"What occupations do the members of school boards follow?" was another inquiry. The occupations were classified according to the system used by the U.S. Census Bureau. In the cities 33 occupations distributed among 8 of the 9 main groups were found. One hundred fifty-seven reports were received concerning 162 board members. Sixty-five of these were following trade pursuits, 43 professional, 27 manufacturing, domestic and personal service 8, transportation 6, clerical pursuits 4, farming and public service 2 each. In the towns we found 7 of the 8 main groups — domestic and personal service group dropping out — with trade, professional pursuits, and manufacturing and mechanical leading as in the cities and the farmer advancing from eighth place to fourth. In the joint corporations we found clerical pursuits eliminated with only 6 groups remaining. Trades and professional pursuits occupied first and second place, but the farmer occupied third place and manufacturing fourth. I think the reasons for these changes are obvious, and no comments need be made here.

Length in Occupation

The time various members have followed present occupations varied from 1 to 65 years, with an average median of 16.73 years. Less than 8 per cent had followed their present occupations for less than 5 years. I think this part of the report shows that in the main school board members have been actively identified with the industrial and professional life of the community.

Education

From the city reports we have the educational records of 146 of the 162 board members. One went no farther than fourth grade; 139 or 84.79 per cent finished the eighth grade; 78 finished high school; 24 finished college; and 3 did 1 year of postgraduate work. One finished a 3-year normal course and 11 took a business course. We have the record of 112 of the 120 members of town school boards. Ninety-nine finished the elementary schools; 44 the high schools; 12 finished the college course; and 4 did postgraduate work of 1 year or more. Four had a 3-year normal course and 6 had a business course. The educational data concerning the members of joint boards revealed the following: The record of 48 of the 60 reported shows that 48 finished the elementary school work; 24 finished 4 years of high school; 4 remained in college 4 years; 2 attended normal school 3 years;

and 3 took business courses. To summarize, the educational qualifications of board members for the 3 types of schools were as follows:

92.4 per cent finished elementary schools.
47.7 per cent remained in high school 4 years.
13 per cent went to college 4 years or longer.
6.5 per cent had a business course.

The Indiana law gives board members almost unlimited power. Persons are not usually given this power because of their profound wisdom concerning school matters. I leave it to those who pay the bills to decide whether board members should not be pruned of some of their power or their qualifications raised. If horse sense is all that is necessary we perhaps have enough of that.

Salaries

Of course the financial remuneration is not sufficient to justify any board member in devoting much time and attention to school matters, the range of salaries being from $25 to $800 in cities, nothing to $100 in towns and joint corporations. The median salary in cities is $89.30, towns $45, joint boards $25, and trustees $600 to $800. About 20 per cent of these board members have served in former years either as teacher, principal, or township trustee.

Term

The range in years of service on school boards runs from 1 to 21 years. The average median is 3.58 years. This is doubtless long enough for a poor member to serve, but it is far too short for a good member. Only a very small per cent of board members are unmarried. All women are either married or are widows. Only 12.8 per cent were reported as having no children.

Church

In cities, 16 per cent were reported as members of no church; 4.5 per cent were Catholics, leaving about 80 per cent as members of Protestant churches. In the towns, 20 per cent are members of no church, while in the joint systems 21.4 per cent belong to no church either Catholic of Protestant.

Politics

In discussing politics it is difficult to draw conclusions. The returns showed that in the 3 types of schools where the city councils were 100 per cent of one party, the school board was either 100 per cent of the same party or were becoming so as rapidly as appointments could be made. This leads one to believe that party lines have been adhered to very closely in school board matters. Even the legislature two years ago recognized that such a condition existed, because a bill was passed and is now law stating that no more than 2 of 3 board members shall belong to the party in power. Like a great many other good intentions of the State legislature, it remains to be seen whether the situation will be helped or whether the party in power will appoint a member of the opposite party who will submit to everything the majority wishes. I predict that this law will help the schools none. But every school man knows after a little experience in the field that partisan politics is not the only kind that is damning the schools of Indiana. It may not even be the worst kind. But I need not go farther in this paper with that.

Work

We shall now turn to the work which the school boards are doing. We shall mention a number of things but cannot enumerate all in this report.

a. Superintendent. The boards select the superintendent, but they have not as yet in many places come to realize his importance in the school system.

b. Teachers. About 60 per cent of the teachers and principals appointed in the 3 types of schools are appointed upon recommendations of the superintendents. In a number of cases the superintendent seems to be ignored by the board, and it is very noticeable that his opinion counts for more in the larger units. There are some places where the superintendent seems to be all-powerful, but in most of the other 40 per cent, the boards make their own selection. This is especially true in the smaller types of schools. It can partly be accounted for by virtue of the fact that the superintendent may not have been giving his attention to this work during the summer. Promotion and assignment of teachers are left to superintendents in at least 75 per cent of cases.

c. Janitors. The same may be said concerning the employment of janitors, except that a much larger percentage are employed without the superintendent's knowing who is employed, or when.

d. Boards Visit Schools. In a few localities school board members are frequent visitors at the schools. Of the 99 boards reported, about 40 per cent did visit schools either singly or as a body during the year. Of all board visits made, 51.3 per cent were in towns and cities ranging from 800 to 5,000. Matters of discipline and social affairs are left almost exclusively to the superintendent, principals, and teachers. Less than 3 per cent of boards mix in these affairs.

e. Curriculum Making of curricula is left almost entirely to the superintendent.

f. Meetings. About 72 per cent of school boards meet regularly. They usually meet in the office of some member of the board. About 50 per cent of city boards meet in the office of superintendent; 26 per cent of town and perhaps 12 per cent of joint boards meet in the superintendent's office. Curbstone meetings are not unknown. Ninety per cent of the superintendents attend the board meetings regularly, but a few stated that they do not attend unless they are called in by the board. This is found to be true even in city systems.

g. Committees. The reports show that about 5 per cent of boards keep standing committees; that 2 per cent of these are found in cities of 10,000 or less. Business matters are usually handled by the board as a whole or certain members of board; the superintendent, or both, are appointed temporarily.

h. Minutes. Four per cent of cities report that the minutes are inaccurately written and poorly kept. The same is reported concerning 12.5 per cent of towns and 20 per cent of joint corporations. Of all minutes in cities, 17 per cent are written by the superintendent or under his supervision and kept in his office. The same may be said of the towns, but the superintendent has nothing to do with the records in the joint corporations reported.

i. Clerks. Forty-three per cent of boards in the 3 types of schools employ no clerk. Seventy-one per cent of cities, reporting on this item, furnish clerks from 3 to 12 months each year. Cities not furnishing any clerk range from 2,500 population to 8,000. Thirty-eight of 40 towns reported this item. Four of this number furnish clerks only when needed. One hires office help full time. Size of towns, 2,000 to 6,000.

j. Programs. Board meetings occur in cities at stated times. They meet weekly in 8 per cent of cities reporting; twice per month in 26 per cent; monthly in 66 per cent. In towns 75 per cent meet monthly; 5 per cent twice per month; and 20 per cent meet irregularly. Joint corporations usually meet monthly.

In cities members are usually punctual. In towns and joint corporations 20 per cent are not. Fifty-five per cent of cities, 20 per cent of towns, and 35 per cent of joint schools have a regular order of procedure. The regular order of business is somewhat as follows:

1. Meeting called to order.
2. Reading of minutes of last meeting.
3. Allowance of bills.
4. Unfinished business.
5. New business.
6. When time permits superintendent may make reports and suggestions.
7. Adjournment more or less irregularly.

Four per cent of cities reported that the superintendent's advice concerning financial and general policies of school is never sought. Eight per cent of towns and 25 per cent of joint corporations report same.

k. Reports. The matter of making reports to school boards concerning the school work and conditions is very generally left to the superintendent, altho in a few cases teachers and principals make such reports directly to board. Frequency of reports varies. It usually depends upon the wish of the superintendent. Items included also vary with conditions. I think types of reports might be improved upon. Too much attention is given to detail. Statistical reports to county and state superintendents are made by the local superintendent in 80 per cent of cities; 67.5 per cent in towns (26.5 per cent with his help ; 40 per cent in joint schools 10 per cent with his help).

The third topic with which the thesis deals is the matter of purchasing supplies. In the cities, 84 per cent of supplies for instruction, 78 per cent of janitor's supplies, 86 per cent of library and laboratory supplies, and 16 per cent of fuel were purchased by superintendent. Sixty-two per cent of fuel was purchased by board and 12 per cent by both board and superintendent. Exclusive of the fuel item, 82.5 per cent of all purchases made, outside of a very few cities that employ a regular purchasing agent, were purchased thru the superintendent's office. We are safe in saying that at least 75 per cent of all school business is done thru the superintendent's office. All other purchases are made by janitors, teachers, supervisors, principals, and board members. In towns the amount purchased by the superintendent was less

He purchased about 75 per cent of supplies for instruction, 50 per cent of janitor's supplies, and 70 per cent of library and laboratory supplies. Ninety-five per cent of fuel was purchased by board. Exclusive of fuel, about 65 per cent of all purchases made were by the superintendent. In the joint schools, about 73 per cent of supplies for instruction were purchased by the superintendent. Twenty-seven per cent of janitor's supplies, 93 per cent of library and laboratory supplies, and 6.6 per cent of fuel were purchased by him. The board purchased about 94 per cent of fuel, 47 per cent of janitor's supplies, and 20 per cent of supplies for instruction.

And thus we might continue at greater length, but enough has already been given to show that a school superintendent in Indiana is largely becoming a school manager. The matter of classroom supervision is being more and more left to supervisors and principals, whether for better or for worse. The natural growth of the school system makes it imperative that the superintendent devote more and more of his time to the business side of school affairs. School boards are by law recognized as doing this work, but the facts show they are not. Most of them cannot take the time to do it if they wished. Besides, there is not enough remuneration to justify their spending their time with such work. I may summarize somewhat as follows:

1. The personnel of school boards consists mainly of men who meet higher qualifications than any Indiana law fixes. They are men of ordinary intelligence. They are connected with the business and professional interests of the community. They have lived long enough in the community in most cases to know something of the educational needs. The age of board members shows them to be men of maturity. The term for a good member is too short three years and six months. They are evidently not free from political influence.

2. The board's duties and those of the superintendent are not clearly marked. Until the special field which each is to occupy is defined there is always a chance that one will interfere with the other.

3. In making of curricula the superintendent has a clear field.

4. The matter of what to report in board meetings and when is deserving of some careful thought and attention.

5. There is need for regular meeting-places and punctuality in attendance.

6. A very large per cent of the business of the schools is being transacted thru the superintendent's office.